Análise de Dados Multivariados

JAMES M. LATTIN
Pós-Graduação na Business School da Stanford University

J. DOUGLAS CARROLL
Pós-Graduação na Escola de Administração da Rutgers University

PAUL E. GREEN
Wharton School, Pennsylvania State University

Revisão Técnica
FLÁVIO SOARES CORDOBA DA SILVA
PhD em Inteligência Artificial pela Edinburgh University,
livre-docente e professor associado do Departamento de
Ciência da Computação no Instituto de Matemática e Estatística da
Universidade de São Paulo (IME-USP)
5

Análise fatorial exploratória

5.1 INTRODUÇÃO

No capítulo anterior, discutimos a análise dos componentes principais, um método para se reduzir a dimensionalidade dos dados multivariados e que facilita o entendimento dos padrões de associação entre as variáveis. Neste capítulo, discutimos a análise fatorial exploratória, um método similar com base em um modelo subjacente diferente, conhecido como modelo de fator comum. Como os métodos são similares, a análise fatorial exploratória é frequentemente usada para atingir os mesmos objetivos da análise dos componentes principais. Nossa meta aqui é enfatizar as diferenças conceituais entre eles. O modelo de fator comum assume pressupostos explícitos sobre como é medida cada variável no conjunto de dados. O modelo sustenta que a variância observada em cada medida pode ser atribuída a um número relativamente pequeno de fatores comuns (isto é, características comuns não observáveis em duas ou mais variáveis) e a um único fator específico (não relacionado com qualquer outro fator subjacente do modelo). Embora possa haver, de fato, mais de um fator específico para uma variável, é impossível distinguir estatisticamente entre um e mais de um fator específico. O objetivo da análise fatorial exploratória é identificar os fatores comuns (separados dos fatores específicos) e explicar sua relação com os dados observados. Portanto, mesmo que o procedimento da solução que usamos para realizar a análise fatorial exploratória seja similar ao dos componentes principais, os modelos subjacentes são diferentes.

Com a análise fatorial exploratória, permitimos que os padrões observados de associação nos dados determinem a solução de fator. No Capítulo 6, retornaremos à análise fatorial confirmatória, na qual começamos com uma noção anterior sobre a estrutura da solução do fator e então testamos para determinar se ela é consistente com os dados. Na análise confirmatória, o modelo subjacente é o mesmo, mas o procedimento de solução é bem diferente do da análise fatorial exploratória.

O modelo de fator comum fornece um quadro de referência explícito que podemos usar para avaliar as propriedades da medida de nossos dados. Nos exemplos fornecidos neste capítulo, vamos supor que a variação em um fator específico seja interpretável puramente como erro de medida. Em geral, para distinguir entre variação específica (isto é, a variação do fator específico) e a variação do erro (isto é, a variação do erro de medida associada com cada variável), é necessário que haja algumas avaliações independentes da confiabilidade da medida (para distinguir-se entre dois fatores específicos separados).
Na ausência dessa avaliação independente, assumimos que a variância no fator específico reflete a quantidade de variância na medida atribuível ao erro. Isso nos dá algum sentido de confiabilidade da medida: quanto menor o erro da variância, mais confiável é a medida. Formalizaremos a noção de confiabilidade no Capítulo 6.

Em nosso desenvolvimento deste capítulo, prestamos particular atenção ao uso da rotação para facilitar a interpretação de uma análise de fator. Como a orientação da solução de fator exploratório é arbitrária, às vezes faz sentido escolher a solução que exiba uma estrutura simples (isto é, aquela que é a mais facilmente interpretável, em um sentido que foi definido inicialmente por Thurstone, 1947). De fato, o uso da rotação é igualmente aplicável quando a análise dos componentes principais (apresentada no Capítulo 4) é utilizada para analisar fatorialmente os dados. Todos os diferentes métodos de rotação desenvolvidos aqui podem ser aplicados à solução dos componentes principais também.

### 5.1.1 Aplicações potenciais

A análise fatorial exploratória pode ser aplicada na maior parte das mesmas aplicações em que se usa a análise de componentes principais; o modelo de fator comum pode ser preferível quando os pressupostos explícitos do modelo de medida são adequados. Relacionadas, a seguir, estão duas aplicações ilustrativas da análise fatorial exploratória: identificação de "traços latentes" ou "características não observáveis" e o uso de escores de fatores na análise de dependência.

Identificando "traços latentes" ou "características não observáveis"

Às vezes, é importante distinguir entre uma variável de dado e o conceito ou ideia que ela se propõe a medir. Quando se está lidando com características físicas, como comprimento ou peso (e um instrumento de medida é altamente preciso), tal distinção pode ser desnecessária porque a propriedade do objeto é quase que perfeitamente observável. No entanto, quando se está lidando com atitudes, crenças, percepções e outras noções psicológicas, nossos instrumentos de medida são, no melhor dos casos, imperfeitos.

Na área de marketing, os pesquisadores podem interessar-se por obter um conceito particular (por exemplo, "satisfação do cliente") para entender-lo melhor e definir como ele é influenciado pelas ações da empresa. Pode ser difícil, se não impossível, projetar uma única pergunta na pesquisa que capte exatamente um constructo como a satisfação do cliente. Em lugar disso, o pesquisador pode elaborar todo um questionário com várias questões, cada uma insuficiente, mas projetada para captar alguma faceta da satisfação do consumidor. Com o uso da análise fatorial, é possível identificar a fonte da variância comum subjacente e, dessa forma, definir o número de fatores subjacentes que explicam a variância em análise subsequente e desenvolvimento do modelo.

Aaker (1997) usou a análise fatorial exploratória para identificar as diferentes dimensões da personalidade de marca. Trabalhando com um conjunto de 114 traços de personalidade (triais de um conjunto de 309 traços possíveis gerados por estudos de personalidade em psicologia e pesquisa de mercado), Aaker pediu aos respondentes que classificassem cada marca de um grupo de 10 marcas em relação a cada um dos 114 traços de personalidade, usando uma escala de cinco pontos (de 1 = absolutamente não descritivo até 5 = extremamente descritivo). Ela utilizou quatro conjuntos diferentes de marcas, cada um contendo uma marca com foco (uma marca de jeanswear Levi’s) e outras nove (todas de marcas nacionais destacadas e bem conhecidas, representando uma ampla variedade de categorias comerciais diferentes), perfazendo um total de 37 marcas diferentes. Aaker calculou a média das classificações dos indivíduos (cada marca foi classificada em relação a cada traço por aproximadamente 150 ou 160 respondentes, com a exceção da Levi’s, que foi classificada por todos) e depois realizou a análise fatorial da matriz de correlação 114 × 114 dos traços.
Tabela 5.2 Solução de dois fatores para os dados de Roberts: matriz de cargas de fator

<table>
<thead>
<tr>
<th></th>
<th>Atraente</th>
<th>Sensato</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luxo</td>
<td>0,884</td>
<td>-0,051</td>
</tr>
<tr>
<td>Estilo</td>
<td>0,748</td>
<td>0,153</td>
</tr>
<tr>
<td>Confiabilidade</td>
<td>0,396</td>
<td>0,691</td>
</tr>
<tr>
<td>Economia de combustível</td>
<td>-0,202</td>
<td>0,786</td>
</tr>
<tr>
<td>Segurança</td>
<td>0,720</td>
<td>0,172</td>
</tr>
<tr>
<td>Manutenção</td>
<td>0,149</td>
<td>0,756</td>
</tr>
<tr>
<td>Qualidade</td>
<td>0,501</td>
<td>0,650</td>
</tr>
<tr>
<td>Durabilidade</td>
<td>0,386</td>
<td>0,677</td>
</tr>
<tr>
<td>Desempenho</td>
<td>0,686</td>
<td>0,391</td>
</tr>
</tbody>
</table>

Modelo de fator parcimonioso, Roberts foi capaz de avaliar a posição relativa do conceito do novo carro (apresentado na Figura 5.1) e prever com exatidão sua participação no mercado.

Figura 5.1 Mapa dos escores fatoriais de automóveis existentes e do conceito do novo carro (Fonte: Roberts, 1984).

5.2 ANÁLISE FATORIAL EXPLORATÓRIA: SEU FUNCIONAMENTO

5.2.1 Intuição
Assim como nos componentes principais, a intuição subjacente à análise fatorial é mais bem explicada com um exemplo simples. Considere o teste psicológico com crianças como o conduzido por Holzinger e Swineford (1939). Foram administrados vários testes diferentes em crianças de 7º e 8º séries (n = 145) e, para simplificar, fuemos os seguintes cinco testes: compreensão de parágrafo (PARA), complementação de sentenças (SENT), significado das palavras (WORD), adição (ADD) e contagem de pontos (DOTS). Usamos as variáveis X₁ a X₅ para representar esses cinco testes diferentes. As correlações entre esses testes (com base em uma amostra de 145 crianças) são apresentadas na Tabela 5.3.

Com a análise fatorial, nosso pressuposto é que os escores observados no teste podem ser descritos como uma função de um pequeno número de fatores comuns subjacentes e um conjunto de fatores específicos (um fator específico para cada teste). Digamos, por exemplo, que acreditamos haver um único fator comum que subjaz os escores de testes dos estudantes, que representamos por ξ (a letra minúscula grega “xi”). Esse fator pode refletir a inteligência ou a capacidade inata de cada estudante de se fazer o teste. Nosso modelo de um fator diz que os escores observados Xᵢ para o teste i são uma função do fator ξ (que é comum a todos os testes) e um fator específico – vamos chamá-lo δᵢ (a letra minúscula grega “delta”) – que é específico do teste i. Em várias abordagens para estimar os parâmetros do modelo de fator comum, tudo que se pode estimar realmente é a variação única (ou a “singularidade”), que é a soma da variação desse fator específico e o erro de medida. Assumiremos que essa variação específica se deve inteiramente ao erro de medida, refletindo a incapacidade do teste de captar perfetamente o fator comum subjacente. Como os fatores específicos afetam somente suas medidas específicas (isto é, o fator δᵢ afeta somente o teste δᵢ), procederemos com a suposição de que esses fatores específicos δᵢ são mutuamente não correlacionados (isto é, a correlação entre δᵢ e δⱼ é igual a zero para todos os i e j) e não correlacionados com o fator comum subjacente ξ. Ambas são suposições padrão do modelo de fator comum.

Figura 5.2 Diagrama de caminhos de um modelo de fator único com cinco variáveis.

O modelo de um fator pode ser representado pelo diagrama da Figura 5.2. As flechas no diagrama que apontam para cada escore do teste (essas são as nossas medidas observadas, representadas pelas caixas) indicam uma fonte de variação que contribui para a medida. Nesse caso, cada medida Xᵢ possui duas fontes que contribuem para a variação (ambas não observáveis): o fator comum ξ e o fator específico δᵢ.

Em forma de equação, podemos escrever o seguinte:

\[X₁ = λ₁ξ + δ₁ \]
\[X₂ = λ₂ξ + δ₂ \]
\[X₃ = λ₃ξ + δ₃ \]
\[X₄ = λ₄ξ + δ₄ \]
\[X₅ = λ₅ξ + δ₅ \]

Nessas equações, os coeficientes λ reflectem o grau com que cada medida X reflete o fator comum subjacente ξ. Pressupondo que tanto X quanto ξ sejam variáveis padronizadas (isto é, com média zero e variação 1), a variação de Xᵢ pode ser decomposta como segue:

\[\text{var}(Xᵢ) = \text{var}(λᵢξ + δᵢ) = λᵢ² + \text{var}(δᵢ) = 1 \]

Como as variáveis são padronizadas, λᵢ é interpretável como um coeficiente de correlação. O termo λᵢ² é interpretável como a proporção da variação em Xᵢ explicada pelo fator comum ξ, e é chamado de comunalidade de Xᵢ. A variação remanescente em Xᵢ é explicada pelo fator específico δᵢ.
Se usarmos \( \theta_0^2 = \text{var}(\delta) \) para representar essa variância do fator específico (que pressupomos refletir o erro de medida em \( X_1 \)), a comunidade de \( X_1 \) é igual a

\[
1 - \theta_0^2
\]

A medida que a comunidade de \( X_1 \) se aproxima de 1 (isto é, à medida que a variância do erro vai a zero), sugere-se que \( X_1 \) é quase uma medida perfeita do fator comum subjacente \( \xi \). Se, por outro lado, \( \xi \) não for de maneira alguma captado pela medida de \( X_1 \) (isso poderia acontecer se \( X_1 \) fosse um teste mal projetado da inteligência do estudante), o coeficiente \( \lambda \) poderia ser próximo de zero e quase todas as variações em \( X_1 \) poderiam ser explicadas pelo fator específico \( \delta \).

Nesta ilustração particular, iniciamos com a suposição de que um único fator comum (como a inteligência geral) \( \xi \) subjacente aos escoros do teste no estudo de Holzinger e Swineford. Entretanto, talvez o desempenho individual no teste seja uma função de mais de uma capacidade subjacente. Por exemplo, poderíamos acreditar que o desempenho individual de um estudante em testes é determinado por duas fontes diferentes de capacidade inata – a aptidão verbal do estudante (que representamos como \( \xi_1 \)) e a aptidão quantitativa (representada por \( \xi_2 \)) – e que esses dois fatores diferentes entram em jogo em diferentes níveis, dependendo do tipo de teste. Esse modelo de dois fatores poderia ser representado como apresentado no diagrama da Figura 5.3.

![Diagrama de Caminhos de um Modelo de Dois Fatores com Cinco Variáveis](image)

**Figura 5.3** Diagrama dos caminhos de um modelo de dois fatores com cinco variáveis.

No modelo de dois fatores, três fontes agora contribuem para a variação observada em cada escoro de medida do teste: \( X_1 \); os dois fatores comuns \( \xi_1 \) e \( \xi_2 \) e um fator específico \( \delta \). Em forma de equação, o modelo de dois fatores poderia ser escrito como

\[
\begin{align*}
X_1 &= \lambda_{11} \xi_1 + \lambda_{12} \xi_2 + \delta_1 \\
X_2 &= \lambda_{21} \xi_1 + \lambda_{22} \xi_2 + \delta_2 \\
X_3 &= \lambda_{31} \xi_1 + \lambda_{32} \xi_2 + \delta_3 \\
X_4 &= \lambda_{41} \xi_1 + \lambda_{42} \xi_2 + \delta_4 \\
X_5 &= \lambda_{51} \xi_1 + \lambda_{52} \xi_2 + \delta_5
\end{align*}
\]

(5.2)

Como anteriormente, o coeficiente \( \lambda \) reflete a extensão com que cada fator comum contribui com a participação de cada escoro de variável no teste. Se \( X \) e \( \xi \) forem padronizados, os parâmetros \( \lambda \) são interpretáveis como coeficientes de correlação. Se, além disso, assumirmos que os fatores comuns subjacentes não são correlacionados (outra suposição padrão do modelo de fator comum), a comunidade de cada medida do teste é dada pela soma das cargas de fator ao quadrado para essa variável; portanto, a comunidade para \( X_1 \) é dada por \( \lambda_{11}^2 + \lambda_{12}^2 \).

Considere o que ocorre, sob o modelo de dois fatores, quando temos um estudante com alta aptidão verbal (alto \( \xi_1 \)) e baixa aptidão quantitativa (baixo \( \xi_2 \)). Podemos esperar que tal estudante se saia bem em testes que exigem mais capacidade verbal do que capacidade quantitativa. Se o desempenho do estudante na tarefa de completar uma sentença (medida por \( X_1 \)) depende somente de sua aptidão verbal, devemos esperar um valor de \( \lambda_{11} \) próximo de 1 e um valor de \( \lambda_{12} \) próximo de zero.

Em um cenário exploratório – aquele em que não há previamente informações fortes a respeito da estrutura do fator subjacente de desempenho do estudante no teste –, desejamos inferir, a partir dos dados, o número apropriado de fatores subjacentes e os valores dos coeficientes nas equações do modelo do fator comum. A abordagem, que detalhamos mais formalmente na próxima seção, é semelhante à dos componentes principais, já que também aqui tentamos extrair um número menor de fatores que irão representar adequadamente a matriz de correlação observada. A diferença em relação ao modelo de fator comum é que devemos agora explicar os fatores específicos \( \delta \) que não estavam presentes nos componentes principais.

**Procedimento para solução**

Da mesma forma que nos componentes principais, o procedimento para solução da análise fatorial exploratória foca a decomposição da covariância ou matriz de correlação de \( X \). A diferença entre as duas abordagens são os fatores específicos que formam o modelo fatorial comum. Como assume-se que esses fatores específicos são mutuamente não correlacionados e independentes dos fatores comuns subjacentes, eles somente contribuem para a diagonal da matriz de covariância. Isso é facilmente observado quando examinamos o elemento diagonal da matriz de covariância de qualquer medida de \( X_i \),

\[
\text{var}(X_i) = \text{var}(\lambda_{i1} \xi_1 + \lambda_{i2} \xi_2 + \delta_i)
\]

(5.3)

Há nove termos de produtos cruzados na expressão para a variação de \( X_i \). Seguindo a partir dos pressupostos de independência impostos pelo modelo de fator comum (e assumindo, além disso, que os fatores comuns foram padronizados com variância igual a 1), todos os termos da covariância abandonam o modelo, resultando em

\[
\text{var}(X_i) = \lambda_{i1}^2 + \lambda_{i2}^2 + \delta_i^2 = 1
\]

(5.4)

onde \( \delta_i^2 = \text{var}(\delta) \). Portanto, o único lugar em que os fatores específicos aparecem na matriz de covariância de \( X \) é na diagonal, onde contribuem com uma parcela da variação de cada medida.

O que aconteceria se conheçéssamos a variação de cada fator específico antecipadamente? Poderíamos, então, subtrair esses valores da diagonal da matriz de covariância. Restaria, nesse caso, uma matriz na qual a única fonte de variação e covariância seria atribuível aos fatores comuns subjacentes. Poderíamos, assim, utilizar uma abordagem de componentes principais para decompor essa matriz e encontrar os fatores comuns. Isso é essencialmente o procedimento de solução que usamos na análise fatorial exploratória. Na análise de componentes principais, decomponhamos a matriz de correlação \( R \) (que tem 1 em sua diagonal), mas, para o modelo de fator comum, decomponhamos uma matriz de correlação com elementos diagonais 1 – \( \theta_0^2 \). Com efeito, começamos subtraindo a variação atribuível aos fatores específicos no modelo (lembrando que tal variação pode ser interpretável como erro de medida ou qualquer outra fonte de variação específica para os testes específicos e não correlacionada a outros fatores no modelo), deixando apenas as comunidades na diagonal. Então, tentamos explicar a variação restante com os fatores comuns. Se soubermos os valores das comunidades, poderemos então solucionar o problema, lançando mão dos mesmos procedimentos que usamos com os componentes principais. Esse processo é chamado, às vezes, de *factoração de eixo principal* (ou simplesmente *abordagem de componentes principais* a modelo de fator comum). A estimativa das comunidades é um aspecto não trivial desse procedimento de solução, e há muitas abordagens diferentes.

Vamos supor, por enquanto, que sabemos alguma coisa sobre o montante da variação atribuível aos fatores específicos em nosso exemplo ilustrativo (mais tarde, nesta seção, discutiremos o que fazer quando esses valores não são conhecidos). Para esse conjunto específico de testes, digamos que aproximadamente metade da variação em cada fator é atribuído ao fator específico – isto é, começamos por estabelecer \( \theta_0^2 = 0.50 \) para todos os \( i \). Naturalmente, não é necessário que as variações do erro sejam iguais em todas as medidas (e, em geral, elas não são); esse é simplesmente um ponto de partida que usamos para ilustrar a aplicação do método. Substituímos os elementos diagonais de \( R \) pelos valores da comunidade 1 – \( \theta_0^2 = 0.50 \) para todos os \( i \) e efetuamos uma decomposição matricial da matriz modificada. Os autovalores são apresentados a seguir:

\[
\lambda_1 = 2.187 \quad \lambda_2 = 1.022 \quad \lambda_3 = -0.135 \quad \lambda_4 = -0.089 \quad \lambda_5 = 0.015
\]
A primeira coisa que observamos é que esses autovalores parecem diferentes daqueles da análise de componentes principais. Nem todos os autovalores são positivos e sua soma não é mais $p$ (o número de variáveis na análise). A razão para isso é que, subtraindo a variação atribuível aos fatores específicos, reduzimos as informações restantes de que necessitamos para determinar os fatores comuns. A variação já explicada pelos fatores específicos no modelo é $0,50 + 0,50 + 0,50 + 0,50 + 0,50 = 2,5$, que é $2,5/5,0 = 50\%$ da variação total nos escores dos cinco testes originais. A soma dos elementos diagonais é também $2,5$, que é o valor remanescente da variação a ser explicada pelos fatores comuns.

Agora, a questão permanece: de quantos fatores comuns necessitamos? Observe que o critério para seleção do número de fatores a serem extraídos modificou-se com relação à análise de componentes principais. Tendo já determinado alguma variação com os fatores específicos, nossa meta agora é inferir a maior variação restante possível com os fatores comuns. Estamos, portanto, buscando todos os autovalores que sejam significativamente diferentes de zero. Nesse caso, há dois, sugerindo que extraíamos $c = 2$ fatores comuns.

A matriz de cargas fatoriais (isto é, a matriz cujos valores são as correlações entre as variáveis originais e os fatores comuns) para a solução de dois fatores é apresentada na Tabela 5.4. O padrão da estrutura do fator mostra que os primeiros três testes (compreensão de parágrafo, compreensão da sentença e significado da palavra) possuem cargas mais altas no primeiro fator comum (com todas as correlações ao redor de 0,8), enquanto os dois últimos testes (adição e contagem de pontos) têm cargas mais altas no segundo fator comum (com correlações ao redor de 0,6). Isso é consistente com a interpretação de que o primeiro fator reflete a aptidão verbal do estudante enquanto o segundo fator indica aptidão quantitativa.

Tabela 5.4 Matriz das cargas fatoriais para um modelo de dois fatores (com comunidades aproximadas)

<table>
<thead>
<tr>
<th></th>
<th>Fator 1</th>
<th>Fator 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARA</td>
<td>0,7722</td>
<td>-0,2351</td>
</tr>
<tr>
<td>SENT</td>
<td>0,7838</td>
<td>-0,1576</td>
</tr>
<tr>
<td>WORD</td>
<td>0,7562</td>
<td>-0,2372</td>
</tr>
<tr>
<td>ADD</td>
<td>0,4293</td>
<td>0,6017</td>
</tr>
<tr>
<td>DOTS</td>
<td>0,3476</td>
<td>0,6506</td>
</tr>
</tbody>
</table>

Utilizando as cargas fatoriais, podemos também calcular a proporção da variação em cada teste explicada pelos dois fatores comuns. Essa proporção é chamada de comunidade. Por exemplo, para o teste de compreensão de parágrafo ($X_1$), os fatores comuns determinam $(0,77)^2 + (-0,24)^2 = 0,65$ ou 65% da variação no teste. O valor dessa comunidade é mais alto que o valor de $1 - \hat{\theta}_i^2 = 0,50$ com o qual iniciamos nossa análise, devido ao fato de nosso ponto de partida ter sido baseado em alguma noção anterior que era somente aproximada. Para uma amostra particular de estudantes, é provável que o valor exato seja diferente.

É possível refinar nossas estimativas iniciais através de um processo iterativo e, assim, finalmente chegar a um resultado consistente. Substituímos nossa estimativa inicial de comunidade pelos resultados da análise fatorial da primeira rodada e efetuamos novamente a análise. Assim, podemos substituir o valor de 0,50 na diagonal para $X_1$ (compreensão de parágrafo) por um novo valor 0,65, e assim por diante para as variáveis restantes. Podemos continuar com esse processo por mais uma iteração, duas iterações ou até que haja convergência – isto é, até que a mudança nas comunidades entre as rodadas seja suficientemente pequena. Esse processo iterativo é frequentemente usado para a estimativa de comunidade. A abordagem resultante é, às vezes, chamada de método de fator principal.

E se não tivéssemos fortes informações preliminares sobre o erro de medida em nossos dados (isto é, a variação não sistemática em nossas medidas não relacionada aos fatores comuns subjacentes)? O que usariamos, então, para uma estimativa inicial de comunidade? Uma medida amplamente utilizada é a correlação múltipla quadrática (SMC – squared multiple correlation), que é o montante de variação em uma variável explicada por todas as outras variáveis no conjunto de dados. Se, por exemplo, quisermos usar a SMC como nossa estimativa inicial de comunidade de $X_1$ (compreensão de parágrafo), podemos fazer a regressão de $X_1$ sobre as variáveis restantes $X_2$, $X_3$, $X_4$ e $X_5$ e usar o valor de $R^2$.

O que torna a SMC uma boa estimativa da comunidade? Lembre-se de que a comunidade é a proporção da variância em X explicada pelos fatores comuns $\xi$. Embora preferíssemos, por causa disso, fazer a regressão de $X$ sobre os fatores comuns $\xi$, temos o problema de que os fatores comuns não serão observáveis. O que temos, no entanto, são as variáveis restantes $X_i$ e cada uma das quais reflete os fatores $\xi$ subjacentes (embora imperfeitamente). Como são medidos com erro, sua capacidade de explicar a variância em $X$ é atenuada. Assim, a SMC serve como um limite inferior à verdadeira comunidade. Em geral, quanto mais confiáveis as medidas (isto é, quanto mais baixa a variância do fator especifico), mais precisa é a SMC como uma estimativa da comunidade.

A Tabela 5.5 mostra os resultados da análise de fator comuns dos dados de escote do teste com o uso da SMC como estimativa inicial de comunidade. Utilizamos um processo iterativo como aquele descrito acima: substituímos a estimativa de comunidade refinada para o valor inicial e continuamos até que haja uma mudança insignificante nos sucessivos valores de comunidade.

Tabela 5.5 Análise fatorial usando SMCs como estimativas iniciais

<table>
<thead>
<tr>
<th>Estimativa Prévia de Comunidade: SMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARA</td>
</tr>
<tr>
<td>SENT</td>
</tr>
<tr>
<td>WORD</td>
</tr>
<tr>
<td>ADD</td>
</tr>
<tr>
<td>DOTS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Autovalores Finais</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2,2826</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Padrão de Fator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fator 1</td>
</tr>
<tr>
<td>PARA</td>
</tr>
<tr>
<td>SENT</td>
</tr>
<tr>
<td>WORD</td>
</tr>
<tr>
<td>ADD</td>
</tr>
<tr>
<td>DOTS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variação Explicada por Cada Fator</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,2826</td>
</tr>
</tbody>
</table>

As diferenças entre as cargas de fator iniciais na Tabela 5.4 e os resultados finais na Tabela 5.5 não são substanciais. O padrão de fator leva à mesma interpretação substantiva dos fatores. A divisão da variação é também quase a mesma: o montante da variação explicado pelos fatores comuns é ainda aproximadamente de 66%. Comparadas com as nossas estimativas iniciais, as comunidades associadas com os dois últimos testes (adição e contagem de pontos) são mais baixas do que com os três primeiros testes: para $X_4$ e $X_5$, as comunidades são aproximadamente 0,60; para $X_1$, $X_2$ e $X_3$, as comunidades são aproximadamente 0,70. Se considerarmos essa fonte de variação um erro de medida, podemos concluir que os dois últimos testes são consideravelmente menos confiáveis do que os três primeiros.

Uma palavra de cautela é necessária em se tratando da abordagem iterativa para estimar comunidades. É se o teste de contagem de pontos for deixado de lado, sobrando a adição como única medida de aptidão quantitativa do estudante? Nessas circunstâncias, nossa estimativa inicial de comunidade...
para a adição, baseada em SMC, será bastante baixa (porque nenhum dos três primeiros testes capta o segundo fator comum). Assim, é mais provável concluirmos que a adição é um teste altamente não confiável quando, de fato, é a única medida de um constructo subjacente importante. É esse problema de estimativa de comunidade que leva alguns pesquisadores a preferir a análise de componentes principais em lugar da análise factorial.

**Rodando a solução factorial**

Como apresentamos na Seção 5.2.2 a seguir, o modelo de fator comum efetivamente possui um número infinito de soluções, cada uma equivalente no que diz respeito a sua capacidade de reproduzir a matriz de covariância observada. A razão para isso é que a orientação da solução factorial (isto é, a escolha dos vetores de base que descrevem o sistema de coordenadas) é, no final das contas, arbitrária. Isso é chamado de indeterminação rotacional do modelo de fator comum. Com a análise de componentes principais, definimos o problema de modo a resolver a indeterminação: o primeiro componente principal é a combinação linear dos dados originais (apropriadamente escalonados) com a maior variância, o segundo componente principal é a combinação linear com a maior variância a seguir (sujeita a ser não correlacionada com a primeira), e assim por diante. Isso assegura uma única (embora arbitrária) solução.

Se a orientação da solução factorial for definitivamente arbitrária, por que não escolher uma que nos ajude a entender-la e interpretá-la melhor? Tentar propor uma interpretação clara dos fatores comuns subjacentes usando a matriz de cargas fatoriais pode ser uma tarefa onerosa e frustrante. Seria vantajoso poder escolher uma orientação diferente, rodando-a a solução factorial, de tal modo que a matriz de cargas fatoriais fosse simplificada. A questão permanece: como escolher tal rotação? Como operacionalizar a meta de simplificar a matriz de cargas fatoriais e depois encontrar uma rotação da solução que nos deixará mais próximos dessa meta?

A abordagem mais popular para se encontrar a matriz T de rotação (por enquanto, nós limitaremos à discussão das rotações ortogonais, que preservam a independência dos fatores comuns subjacentes) baseia-se nos princípios da estrutura simplificada desenvolvida por Thurstone (1947). Thurstone acreditava que a maioria dos domínios de conteúdo provavelmente envolveria diversos fatores latentes (isto é, subjacentes e não observados). Ele também supôs que qualquer variável observada isoladamente poderia estar correlacionada a apenas uns poucos fatores subjacentes, e que qualquer fator individualmente poderia estar associado com apenas algumas variáveis. A ideia geral, então, era encontrar agrupamentos de variáveis na medida do possível, com cada um definindo somente um fator. De modo mais generalizado, gostaríamos de encontrar uma orientação dos eixos do fator de tal modo que cada teste (ou outra variável) tenha cargas relativamente altas (positivas ou negativas) em apenas alguns fatores, sendo a maioria das outras cargas próximas de zero.

Se o raciocínio subjacente da estrutura simples for válido, a matriz de cargas fatoriais deve exibir um tipo particular de padrão (Comrey, 1973):

1. A maioria das cargas de qualquer fator específico (coluna) deve ser pequena (tão próxima de zero quanto possível), e somente algumas cargas devem possuir valor absoluto elevado.
2. Uma linha específica da matriz de cargas, contendo as cargas de uma dada variável com cada fator, deve exibir cargas diferentes de zero em apenas um ou não mais que em alguns poucos fatores.
3. Qualquer par de fatores (colunas) deve exibir diferentes padrões de cargas. De outra forma, não se poderia distinguir os dois fatores representados por essas colunas.

Um exemplo hipotético que demonstra a noção de estrutura simples é apresentado na Tabela 5.6 e na Figura 5.4. Imagine um estudo das percepções do consumidor de analgésicos, no qual cada sujeito é solicitado a classificar sua marca preferida de acordo com seis atributos:

1. Não indispõe o estômago.
2. Não possui efeitos colaterais indesejáveis.
3. Acaba com a dor.
4. Funciona rapidamente.

<table>
<thead>
<tr>
<th>Atributo</th>
<th>Solução não rodada</th>
<th>Solução rodada</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fator 1</td>
<td>Fator 2</td>
</tr>
<tr>
<td>Não indispõe o estômago</td>
<td>0,579</td>
<td>-0,452</td>
</tr>
<tr>
<td>Sem efeitos colaterais indesejáveis</td>
<td>0,522</td>
<td>-0,572</td>
</tr>
<tr>
<td>Acaba com a dor</td>
<td>0,645</td>
<td>0,436</td>
</tr>
<tr>
<td>Funciona rapidamente</td>
<td>0,542</td>
<td>0,542</td>
</tr>
<tr>
<td>Mantém-me acordado</td>
<td>-0,476</td>
<td>0,596</td>
</tr>
<tr>
<td>Alívio limitado</td>
<td>-0,613</td>
<td>-0,439</td>
</tr>
</tbody>
</table>

Variância explicada por

<table>
<thead>
<tr>
<th>Solução não rodada</th>
<th>Solução rodada</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,921</td>
<td>1,562</td>
</tr>
<tr>
<td>1,765</td>
<td>1,718</td>
</tr>
</tbody>
</table>

**Figura 5.4** Gráfico das cargas fatoriais para atributos dos analgésicos: rodados e não rodados.

A matriz de cargas fatoriais na primeira parte da Tabela 5.6 mostra a solução não rodada de uma análise factorial com dois fatores comuns. (Lembre-se de que esses fatores não rodados são orientados de tal forma que o primeiro fator comum explique o valor máximo de variação; o segundo fator comum explica tanta variação residual quanto possível, e pode não estar correlacionado com o primeiro fator comum.) Observe que todos os 12 elementos da matriz de cargas fatoriais são relativamente altos (todos maiores que 0,4 em valor absoluto). Com tantas cargas cruzadas substanciais, é difícil obter uma interpretação clara dos fatores.

A Figura 5.4 apresenta as cargas fatoriais e mostra como é possível rodar o espaço fatorial de modo que os atributos se alinhem mais proximamente aos fatores comuns. A ideia é escolher o ângulo de rotação para que as projeções de cada atributo sobre os fatores comuns sejam ou elevados (isto é, próximos a 1 em valor absoluto) ou reduzidos (próximos a zero). Isso reduz essencialmente as cargas cruzadas e desloca a matriz de cargas na direção de uma estrutura simples.

A matriz de cargas fatoriais rodada na Tabela 5.6 agora exibe algo próximo a uma estrutura simples. Os atributos 3 e 4 ("Acaba com a dor" e "Funciona rapidamente") carregam-se positivamente no primeiro fator comum, e o atributo 6 ("Fornec. alívio limitado") carrega-se negativamente. Os atributos
5.2.2 Mecânica

Nesse tratamento mais formal do modelo de análise fatorial, é interessante retornarmos ao nosso desenvolvimento do modelo de componentes principais do Capítulo 4. Mostramos que resolver o problema de componentes principais era equivalente a fazer uma decomposição do valor único da matriz $X$ de dados padronizados, como apresentado a seguir:

$$X = Z\Lambda^{1/2}U'$$  \hspace{1cm} (5.5)

onde $Z$ representa os componentes principais padronizados (todos mutualmente não correlacionados), $\Lambda^{1/2}$ é uma matriz diagonal com os desvios-padrão dos componentes principais definindo a diagonal e $U$ é a matriz de autovalores (todos mutualmente ortogonais). Utilizando essa descoberta, podemos reescrever a matriz de correlação da amostra $R$ como uma função dos autovalores e autovetores da decomposição do valor único. A matriz de correlação é dada por

$$R = 1/(n-1)XX'$$  \hspace{1cm} (5.6)

Substituindo-se a decomposição de valor único da Equação (5.5) na Equação (5.6) e simplificando, temos

$$R = 1/(n-1)(Z\Lambda^{1/2}U')'Z\Lambda^{1/2}U'$$

$$= 1/(n-1)(U\Lambda^{1/2}Z)'(U\Lambda^{1/2}Z)$$  \hspace{1cm} (5.7)

porque $1/(n-1)Z'Z$ é apenas uma matriz identidade. Se, além disso, lembrarmos que o produto da matriz $U\Lambda^{1/2}$ é apenas a matriz $F$ de cargas fatoriais (isto é, a matriz de correlação entre a matriz original de dados $X$ e a matriz de componentes principais $Z$), podemos simplificar a reexpressão de $R$ como segue:

$$R = FF'$$  \hspace{1cm} (5.8)

Uma vez que nosso objetivo com a análise de componentes principais é, com frequência, obter alguma forma de redução de dimensão, tentamos extrair um subconjunto de componentes $c$ (onde $c < p$, o número de variáveis de $X$) que se aproxime muito de $R$. Assim, nos componentes principais, podemos determinar

$$R \approx F_cF'_c$$  \hspace{1cm} (5.9)

onde $F_c$ representa somente as $c$ primeiras colunas da matriz $F$ de cargas fatoriais.

Em análise fatorial exploratória, também tentamos aproximar a matriz de correlação $R$, mas usando um modelo diferente. Em lugar de substituir a decomposição de valor único para $X$ na Equação (5.6), usamos o modelo de fator comum esboçado anteriormente. O modelo geral com $c$ fatores comuns pode ser escrito como segue:

$$X_1 = \lambda_1\xi_1 + \lambda_2\xi_2 + \cdots + \lambda_c\xi_c + \delta_1$$

$$X_2 = \lambda_1\xi_1 + \lambda_2\xi_2 + \cdots + \lambda_c\xi_c + \delta_2$$

$$X_3 = \lambda_1\xi_1 + \lambda_2\xi_2 + \cdots + \lambda_c\xi_c + \delta_3$$

$$\vdots$$

$$X_p = \lambda_1\xi_1 + \lambda_2\xi_2 + \cdots + \lambda_c\xi_c + \delta_p$$

Na notação matricial, o modelo de fator comum torna-se

$$X = \Xi\Lambda_c' + \Delta$$  \hspace{1cm} (5.11)

onde $\Xi = (\xi_1, \xi_2, \ldots, \xi_c)$ e $\Lambda = (\lambda_1, \lambda_2, \ldots, \lambda_c)$ é uma matriz de coeficientes $p \times c$. Além disso, fazemos os seguintes três pressupostos sobre os componentes do modelo de fator comum.

**PRESSUPPOSTOS PARA O MODELO DE FATOR COMUM**

1. Os fatores comuns $\xi$ são mutuamente não correlacionados com variação unitária:

$$\frac{1}{(n-1)}\Xi'\Xi = I$$

2. Os fatores específicos $\delta$ são mutuamente não correlacionados com a matriz de covariância diagonal:

$$\Theta = \frac{1}{(n-1)}\Delta'\Delta = \text{diag}(\delta_{11}, \delta_{22}, \ldots, \delta_{pp})$$

3. Os fatores comuns $\xi$ e os fatores específicos $\delta$ não são correlacionados:

$$\Xi'\Delta = 0$$

Substituímos agora o modelo de fator comum da Equação (5.11) na Equação (5.6) para obter nossa aproximação à matriz de correlação $R$:

$$R = 1/(n-1)((\Xi\Lambda_c' + \Delta)(\Xi\Lambda_c' + \Delta)'$$

$$= 1/(n-1)((\Xi\Lambda_c' + \Delta)(\Xi\Lambda_c' + \Delta))$$  \hspace{1cm} (5.12)

Pelo pressuposto 3 do modelo de fator comum, o segundo e o terceiro termos na expressão entre parênteses acima vão para zero. Pelo pressuposto 1, podemos substituir a expressão $1/(n-1)\Xi\Xi'$ do primeiro termo pela matriz identidade $I$. E, pelo pressuposto 2, o último termo na expressão se torna $\Theta$. Fazendo essas simplificações, chegamos a

$$R = \Lambda_c\Lambda_c' + \Theta$$  \hspace{1cm} (5.13)

ou

$$R - \Theta = \Lambda_c\Lambda_c'$$  \hspace{1cm} (5.14)

Uma comparação da Equação (5.9) do modelo de componentes principais com a Equação (5.14) do modelo de fator comum revela a semelhança entre as duas abordagens — e sua diferença essencial. Em cada caso, fazemos uma decomposição de matriz de uma matriz simetricamente quadrada. A matriz $\Lambda_c$ é exatamente análoga à matriz $F_c$, é uma matriz de cargas fatoriais, cujos elementos são interpretáveis como as correlações entre as variáveis originais $X$ e os fatores comuns extraídos $c$. 
Indeterminação rotacional

Nos componentes principais, escolhemos cada componente de modo sequencial para explicar o valor máximo possível de variação em nossos dados originais, com restrição de que haverá correlação com todos os componentes previamente selecionados. Isso garantia uma solução única (embora alguma arbitria em termos de orientação escolhida). Com o modelo de fator comum, no entanto, não impusemos tal restrição. Como consequência, há efetivamente um número infinito de soluções que são idênticas no seu grau de aproximação da matriz $R - \Theta$. Referimos-nos a essa propriedade como indeterminação rotacional do modelo de fator comum.

Demonstraremos inicialmente essa propriedade com um exemplo. Considere a solução de dois fatores para os dados de escore de teste apresentados na Tabela 5.5. Um gráfico das cargas fatoriais da tabela é mostrado na Figura 5.5. Mudamos agora a orientação dos fatores rodando-os 30 graus em sentido horário. A rotação (que é executada pela multiplicação da matriz) preserva a ortogonalidade dos dois fatores. Do Capítulo 2, sabemos que uma matriz, para realizar uma rotação ortogonal (em duas dimensões), toma a seguinte forma:

$$
T = \begin{bmatrix}
\cos \alpha & -\sin \alpha \\
\sin \alpha & \cos \alpha 
\end{bmatrix}
$$

(5.15)

Figura 5.5 Diagrama mostrando uma rotação no sentido horário de 30 graus.

Quando o ângulo de rotação $\alpha = -30$ graus, obtemos a seguinte matriz $T$ de rotação ortogonal:

$$
T = \begin{bmatrix}
0.866 & 0.500 \\
-0.500 & 0.866
\end{bmatrix}
$$

(5.16)

Mudando a orientação dos eixos que representam os fatores comuns, também mudamos os valores das cargas fatoriais. Podemos calcular as novas cargas (representadas por $\Lambda'_c$), que são apenas as correlações entre os fatores rodados ($\Xi T$) e as variáveis originais $X$ (dadas por $\Xi \Lambda'_c + A$), simplesmente como

$$
\Lambda'_c = \frac{1}{n-1}(\Xi \Lambda'_c + A)' \Xi T
$$

ou

$$
\Lambda'_c = A'T
$$

porque $1/(n-1)\Xi'\Xi = I$ e $\Lambda'\Xi = I$. As cargas fatoriais rodadas $\Lambda'_c$ são apresentadas na Tabela 5.7 ao lado de cargas originais não rodadas. Como antes, as principais cargas na matriz (isto é, aquelas com os valores absolutos maiores) não se modificaram dramaticamente: os três primeiros testes carregam sobre o primeiro fator e os dois últimos carregam no segundo. Mas observe que as cargas cruzadas na solução rodada mudaram. Os três primeiros testes agora carregam positivamente sobre o segundo fator (ao invés de negativamente), e os dois últimos testes carregam quase exclusivamente sobre o segundo fator (em vez de positivamente sobre os dois fatores).

Tabela 5.7 Cargas fatoriais iniciais e cargas fatoriais rodadas - 30 graus

<table>
<thead>
<tr>
<th></th>
<th>Fator 1</th>
<th>Fator 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARA</td>
<td>0,811</td>
<td>-0,200</td>
</tr>
<tr>
<td>SENT</td>
<td>0,811</td>
<td>-0,106</td>
</tr>
<tr>
<td>WORD</td>
<td>0,779</td>
<td>-0,195</td>
</tr>
<tr>
<td>ADD</td>
<td>0,375</td>
<td>0,586</td>
</tr>
<tr>
<td>DOTS</td>
<td>0,295</td>
<td>0,614</td>
</tr>
<tr>
<td></td>
<td>0,801</td>
<td>0,232</td>
</tr>
<tr>
<td></td>
<td>0,754</td>
<td>0,313</td>
</tr>
<tr>
<td></td>
<td>0,775</td>
<td>0,221</td>
</tr>
<tr>
<td></td>
<td>0,032</td>
<td>0,695</td>
</tr>
<tr>
<td></td>
<td>-0,052</td>
<td>0,679</td>
</tr>
</tbody>
</table>

(5.17)

Multiplicando as matrizes, fica fácil ver que $TT' = I$:

$$
\begin{bmatrix}
\cos^2 \alpha + \sin^2 \alpha &=& 1 \\
\sin \alpha \cos \alpha + \sin \alpha \cos \alpha &=& 0
\end{bmatrix}
$$

(5.17)

Mais genericamente, $TT' = I$ é sempre verdadeiro para uma rotação ortogonal, em qualquer número de dimensões, pois a matriz transposta age no sentido de rodar os eixos de volta na direção oposta, deixando-os com a orientação inicial.

O resultado importante é que as comunidades, a variância explicada pelos fatores comuns e a adequação do modelo de fator comum não são afetadas pela rotação ortogonal — será útil mais tarde, quando consideraremos os modos de melhorar a capacidade de interpretação da solução fatorial.

Rotação fatorial

Para encontrar a rotação apropriada, devemos determinar um meio de quantificar o que queremos dizer com estrutura simples na forma de uma função objetivo. Então, procuraríamos através de todos os ângulos possíveis de rotação e selecionaríamos a matriz de rotação $T$, de tal forma que a matriz de cargas fatoriais rodada $A = \Lambda T$ exiba o maior valor da função objetivo para uma estrutura simples. Há muitas maneiras diferentes de se quantificar a estrutura simples, e cada uma envolve uma forma ligeiramente diferente de rotação. No restante desta seção, discutiremos apenas duas: a rotação varimax de Kaiser e a rotação quartimax.

Rotação varimax de Kaiser. Lembre-se de que cada elemento da matriz de cargas rodada $a_{ik}$ pode ser entendido como a correlação entre a variável $i$ e o fator comum $k$. A carga ao quadrado $a_{ik}^2$ é a proporção da variação na variável $i$ atribuível ao fator comum $k$. Já que escolhemos os fatores comuns como não correlacionados, o valor total da variação explicada por todos os fatores comuns — que denominamos de
comunalidade – é dado pela soma das cargas ao quadrado, ou $h_i^2 = \Sigma_{k} a_{ik}^2$. Para obter uma estrutura simples, gostaríamos de encontrar uma rotação que determine que as cargas ao quadrado $a_{ik}^2$ sejam próximas a 1 ou zero.

O procedimento varimax tenta fazer isso ficando-se nas colunas de $A$: ele escolhe a matriz de rotação $T$ para maximizar a variância total da coluna $a_{ik}^2$. A $k$-ésima variância de coluna é dada pela expressão

$$V_k = \frac{1}{p} \sum_{i=1}^{p} a_{ik}^2 - \frac{1}{p^2} \left( \sum_{i=1}^{p} a_{ik}^2 \right)^2$$

(5.18)

Maximizar a soma dessas variâncias de coluna $V_k$ para todos os fatores $k$ equivale a maximizar a seguinte expressão:

$$V = \sum_{k=1}^{p} \left( \frac{1}{p} \sum_{i=1}^{p} a_{ik}^2 - \frac{1}{p^2} \left( \sum_{i=1}^{p} a_{ik}^2 \right)^2 \right)$$

(5.19)

Observe que a variância máxima é alcançada quando os valores de $a_{ik}^2$ são levados a zero ou 1; por definição, quando o valor de $a_{ik}^2$ se aproxima de 1 para algum fator $k$, todos os outros elementos naquela linha da matriz são levados a tender a zero.

É também possível construir a função objetiva para uma rotação varimax utilizando-se as cargas ao quadrado normalizadas, $a_{ik}^2/h_i^2$. Quando as cargas são normalizadas dessa maneira, o elemento $a_{ik}^2/h_i^2$ é interpretável como a proporção da variância comun na variável $i$ atribuível ao fator comum $k$. Usar essa normalização garante que todas as variáveis recebem peso igual na escolha da rotação (que pode não ser o caso para cargas ao quadrado não normalizadas, quando algumas variáveis possuem baixas comunalidades).

**Rotação quartimax.** Em contraste com a varimax, que foca as colunas da matriz de cargas fatoriais rodadas $A$, a rotação quartimax foca as linhas. A função objetiva para a quartimax depende de as comunalidades de uma variável não serem alteradas pela rotação ortogonal; assim, a expressão $\Sigma_{k} a_{ik}^2$ é constante, independente da matriz de rotação $T$. É também verdade que a soma das comunalidades ao quadrado por todas as variáveis $\Sigma_{k} (\Sigma_{k=1}^{p} a_{ik}^2)^2$ é igualmente uma constante. Expansando essa expressão, o resultado é

$$\sum_{i=1}^{p} \sum_{k=1}^{p} a_{ik}^2 + \sum_{i=1}^{p} \sum_{k=1}^{p} \left( \sum_{k=1}^{p} a_{ik}^2 \right)^2$$

(5.20)

O segundo termo na expressão acima é simplesmente a soma dos produtos cruzados das cargas ao quadrado. Quando uma matriz exibe uma estrutura simples, esse produto deve ser tão pequeno quanto possível (isto é, quando uma variável carrega-se altamente sobre algum fator $k$, a carga deve ser próxima de zero para todos os outros fatores $j$). Uma vez que a expressão anterior é uma constante para todas as matrizes de rotação $T$, maximizar o seu primeiro termo é um modo de se garantir que os termos do produto cruzado sejam pequenos. Portanto, a quartimax escolhe uma matriz de rotação ortogonal $T$ para maximizar

$$Q = \sum_{i=1}^{p} \sum_{k=1}^{p} a_{ik}^2$$

(5.21)

Como na varimax, é possível normalizar as cargas ao quadrado antes de realizar a rotação, dividindo-se pela comunalidade de cada variável.

Levantamos a questão da rotação oblíqua (em que a ortogonalidade mútua dos fatores não é preservada pela rotação) na Seção 5.4.1 a seguir.

**Escores fatoriais.**

Com frequência, a análise fatorial não é um fim em si mesma, mas um passo intermediário no caminho de uma análise posterior dos dados. Para realizar uma análise subsequente, precisamos da localização de cada uma das observações originais no reduzido espaço fatorial. Esses valores são chamados de **escores fatoriais**.

Obter escores fatoriais de um modelo de fator comum não é tão fácil quanto obter escores da análise de componentes principais. Lembre-se de que os escores dos componentes principais são combinações lineares de variáveis originais que podem ser calculadas diretamente, utilizando-se os coeficientes dos autovaleurs da matriz de correlação. Na análise de fator comum, os escores não podem ser calculados exatamente devido à indeterminação introduzida pelos fatores específicos. Em outras palavras, com o modelo de fator comum, não podemos determinar $\xi$ como uma função de $X$ sem conhecer $\Omega$.

Portanto, é necessário calcular os coeficientes dos escores fatoriais que usaremos para calcular os escores fatoriais. Usamos a seguinte expressão linear para aproximar $\Xi$:

$$\Xi = X \Omega$$

(5.22)

onde $B$ é a matriz de coeficientes de escor de fator. Como os valores de $\Xi$ não são diretamente observados, não podemos escolher $B$ usando a regressão dos mínimos quadrados. No entanto, se pré-multiplicarmos cada lado da equação acima por $1/(n - 1)X'_s$, teremos

$$\frac{1}{(n - 1)} X'_s \Xi = \frac{1}{(n - 1)} X'_s X B$$

(5.23)

ou

$$A'_c = RB$$

(5.24)

Podemos garantir que essa igualdade se mantém caso pré-multiplicamos ambos os lados dessa equação por $R^{-1}$, levando à expressão

$$B = R^{-1} A'_c$$

(5.25)

para os coeficientes do escor fatorial. Substituindo essa expressão de $B$ na Equação (5.22), teremos a seguinte expressão para os escores fatoriais estimados:

$$\Xi = X R^{-1} A'_c$$

(5.26)

Observe que, como a matriz de cargas $A'_c$ está sujeita à indeterminação rotacional, os escores de fator identificados na Equação (5.25) não são únicos e são sujeitos à mesma indeterminação rotacional. O que não varia é o produto $\hat{X}_c = A'_c X$, em que $\hat{X}_c$ é a parte apropriada de X explicada pelos fatores comuns $c$.

5.3 **EXEMPLO DE PROBLEMA: PERCEPÇÕES DE CEREAIS PRONTOS PARA COMER**

5.3.1 **DADOS**

Como parte de um estudo das impressões do consumidor sobre cereais prontos para comer, patrocinado pela Kellogg da Austrália, Roberts e Lattin (1991) pesquisaram a percepção desses consumidores em relação às suas marcas favoritas de cereais. Solicitou-se a cada respondente que avaliasse suas três marcas preferidas tendo em mente cada um dos 25 diferentes atributos. Uma escala de cinco pontos foi utilizada para indicar a extensão de cada atributo em relação a cada marca.

**EXEMPLO**

*Criação de um mapa de percepção das marcas de cereais prontos para comer no mercado australiano (Roberts e Lattin, 1991)*

RTF_CEREAL [RTF = Pronto para comer]

Para os propósitos dessa ilustração, selecionamos um subconjunto de dados coletados por Roberts e Lattin, refletindo as avaliações das 12 marcas de cereais mais frequentemente citadas na amostra (no estado original, um total de 40 marcas diferentes foram avaliadas por 121 respondentes, mas apenas um pequeno número de consumidores classificou a maior parte das marcas). Os 25 atributos e as 12 marcas estão relacionados na Tabela 5.8. No total, 116 respondentes forneceram 235 observações sobre as 12 marcas selecionadas.
Tabela 5.8 Lista de 25 atributos de cereais RTE e 12 marcas de cereais RTE

<table>
<thead>
<tr>
<th>Cereais</th>
<th>Atributos</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. All Bran</td>
<td>Satisfaz</td>
</tr>
<tr>
<td>2. Cereal Muesli</td>
<td>Natural</td>
</tr>
<tr>
<td>3. Just Right</td>
<td>Fibra</td>
</tr>
<tr>
<td>4. Kellog's Corn Flakes</td>
<td>Doce</td>
</tr>
<tr>
<td>5. Komplett</td>
<td>Fácil</td>
</tr>
<tr>
<td>6. NutriGrain</td>
<td>Sal</td>
</tr>
<tr>
<td>7. Purina Muesli</td>
<td>Gratificante</td>
</tr>
<tr>
<td>8. Rice Bubbles</td>
<td>Energia</td>
</tr>
<tr>
<td>9. Special K</td>
<td>Divertido</td>
</tr>
<tr>
<td>10. Sustain</td>
<td>Crianças</td>
</tr>
<tr>
<td>11. Vitabrit</td>
<td>Encharcado</td>
</tr>
<tr>
<td>12. Weetbix</td>
<td>Econômico</td>
</tr>
</tbody>
</table>

Um objetivo do estudo era caracterizar o comportamento do consumidor de cereal (isto é, explicar quais marcas o consumidor poderia pensar em adquirir e quais delas sequer consideraria) como uma função das características subjacentes das marcas disponíveis. Para alcançar tal objetivo em nível individual, é necessário um modelo relativamente parcimonioso. Roberts e Lattin, portanto, propuseram usar a análise fatorial para reduzir a dimensionalidade dos dados de atributo e revelar um número menor de fatores subjacentes que explicassem um valor importante da variância nas medidas originais. Vamos agora analisar nosso subconjunto escolhido dos dados de Roberts e Lattin, usando análise fatorial exploratória, e discutir os resultados.

5.3.2 Resultados

Começamos examinando um gráfico scree dos autovalores da análise de componentes principais para determinar o número apropriado de fatores para o modelo de fator comum. O gráfico scree, apresentado na Figura 5.6, sugere que extração de quatro fatores é justificável; observe que três delas parecem formar um “ángulo” no valor da variância explicada após o quarto autovalor. Usando um critério de proporcionalidade (isto é, cada fator comum deve explicar pelo menos tanta variância quanto uma das variáveis originais da análise, o que é diretamente análogo à lógica subjacente da regra de Kaiser para reter fatores na análise de componentes principais), a inclusão do quinto fator é, no melhor dos casos, marginal.

Em seguida, executamos o modelo de fator comum, usando a SMC como nossas estimativas iniciais das comunalidades. As SMC, apresentadas na Tabela 5.9, variam em valor, de um baixo 0,23 (para o

![Gráfico scree para os dados de cereal RTE.](image)

A solução de quatro fatores, incluindo cargas fatoriais não rodadas, é apresentada na Tabela 5.10 a seguir. Observe que alguns dos atributos têm valores relativamente baixos de comunalidade. Por exemplo, o atributo “Fácil” possui um valor final de comunalidade de 0,17 na solução de quatro fatores. Deste modo, temos que 83% da variação nessa medida particular são únicos para o fator. Isso é potencialmente uma preocupação; por essa variável particular, não estamos captando quase nenhuma de suas informações na solução de fator comum.

Há várias razões potenciais para que isso ocorra. Uma pode ser o fato de termos optado por extrair muito poucos fatores e, como consequência, há uma quantidade limitada de informações que podemos explicar nesta medida particular. Em geral, não é uma ideia tentar extrair um fator adicional para determinar se ele altera os resultados (isto é, a comunalidade do atributo aumenta?). Em caso positivo, a capacidade de explicar a variação adicional em alguma medida importante pode justificar a complexidade de uma solução com número mais elevado de fatores.

Uma segunda razão potencial tem a ver com o problema do cálculo da comunalidade mencionada anteriormente. Se “Fácil” é o único atributo que capta a facilidade de uso subjacente do produto, nossa estimativa inicial da comunalidade desse atributo (baseada na SMC) será então baixa. Sem ao menos outra variável projetada para medir o mesmo fator comum (ou alguma avaliação independente da confiabilidade do atributo), não podemos averiguar a verdadeira importância do atributo na análise.

Por fim, uma comunalidade baixa pode também ser tomada como indicação de uma medida ruim (isto é, uma medida na qual há muito considerável associado a erro de medida). Nesse caso particular, parece que “Fácil” é um atributo que se aplica igualmente bem a todos os itens na categoria do produto (afinal de contas, qualquer cereal pronto para consumo deve ser fácil de preparar). Dentro da categoria, pode ser muito difícil para os consumidores distinguir entre as marcas de cereais com base na facilidade de uso. Se esse for o caso, é de se esperar que a maior parte da variância nesse atributo seja não sistemática; como consequência, a solução fatorial teria um componente de variância específica relativamente alta. Em geral, alguma forma de pré-testagem para se estabelecer a propriedade da medida em um dado contexto de estudo ajudará a evitar o problema.

Agora, usamos a rotação como tentativa de facilitar a interpretação da solução de quatro fatores para cereais RTE. A Tabela 5.11 mostra a solução de fator rodado varimax (a solução quartenmax seria muito similar). Para tornar mais fácil a visualização da estrutura nas cargas fatoriais, sublinhamos todos os elementos maiores que 0,40 em cada matriz (se nenhum elemento na fila exceder 0,40, marcamos o elementos de valor mais alto com um sublinhado com pontos).

Voltamos agora à interpretação. Nosso objetivo ao nominar cada fator é sugerir um termo que descreva o domínio do conteúdo das variáveis que carregam altamente naquele fator. Dar nome a um fator é quase sempre uma forma de arte em si. Em geral, é muito pobre utilizar-se apenas dos nomes das variáveis para denominar os fatores, embora, algumas vezes, isso seja inevitável.
Tabela 5.10 Solução factorial não rodada para dados de cereal RTE

<table>
<thead>
<tr>
<th></th>
<th>Fator 1</th>
<th>Fator 2</th>
<th>Fator 3</th>
<th>Fator 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaz</td>
<td>0,7276</td>
<td>0,1030</td>
<td>-0,0674</td>
<td>0,1960</td>
</tr>
<tr>
<td>Natural</td>
<td>0,7323</td>
<td>-0,2422</td>
<td>-0,1077</td>
<td>0,1052</td>
</tr>
<tr>
<td>Fibra</td>
<td>0,7226</td>
<td>-0,2385</td>
<td>-0,3098</td>
<td>0,1607</td>
</tr>
<tr>
<td>Doce</td>
<td>0,0850</td>
<td>0,7407</td>
<td>-0,2053</td>
<td>0,1510</td>
</tr>
<tr>
<td>Fácil</td>
<td>0,3181</td>
<td>0,1454</td>
<td>0,2064</td>
<td>0,1074</td>
</tr>
<tr>
<td>Sal</td>
<td>-0,2062</td>
<td>0,4995</td>
<td>-0,1385</td>
<td>0,3971</td>
</tr>
<tr>
<td>Gratificante</td>
<td>0,7207</td>
<td>0,1809</td>
<td>0,1627</td>
<td>0,1704</td>
</tr>
<tr>
<td>Energia</td>
<td>0,7022</td>
<td>0,1328</td>
<td>-0,0657</td>
<td>0,1264</td>
</tr>
<tr>
<td>Divertido</td>
<td>0,3886</td>
<td>0,4959</td>
<td>0,2101</td>
<td>-0,1579</td>
</tr>
<tr>
<td>Crianças</td>
<td>0,2133</td>
<td>0,2759</td>
<td>0,7335</td>
<td>0,1118</td>
</tr>
<tr>
<td>Encharcado</td>
<td>-0,0984</td>
<td>-0,2319</td>
<td>0,1515</td>
<td>0,4257</td>
</tr>
<tr>
<td>Econômico</td>
<td>0,1508</td>
<td>-0,2361</td>
<td>0,4890</td>
<td>0,1002</td>
</tr>
<tr>
<td>Saúde</td>
<td>0,8081</td>
<td>-0,3075</td>
<td>-0,1068</td>
<td>0,0713</td>
</tr>
<tr>
<td>Família</td>
<td>0,3074</td>
<td>0,2179</td>
<td>0,6731</td>
<td>0,0271</td>
</tr>
<tr>
<td>Calorias</td>
<td>-0,1603</td>
<td>0,5663</td>
<td>-0,1721</td>
<td>0,2151</td>
</tr>
<tr>
<td>Simples</td>
<td>-0,3055</td>
<td>-0,3560</td>
<td>0,2178</td>
<td>0,4086</td>
</tr>
<tr>
<td>Crocante</td>
<td>0,2888</td>
<td>0,4544</td>
<td>0,2143</td>
<td>-0,1927</td>
</tr>
<tr>
<td>Regular</td>
<td>0,5943</td>
<td>-0,1354</td>
<td>-0,1927</td>
<td>0,0682</td>
</tr>
<tr>
<td>Açúcar</td>
<td>-0,2476</td>
<td>0,7213</td>
<td>-0,2467</td>
<td>0,2416</td>
</tr>
<tr>
<td>Fruta</td>
<td>0,3733</td>
<td>0,2534</td>
<td>-0,4802</td>
<td>-0,1569</td>
</tr>
<tr>
<td>Processo</td>
<td>-0,3033</td>
<td>0,2721</td>
<td>0,0030</td>
<td>0,2470</td>
</tr>
<tr>
<td>Qualidade</td>
<td>0,7324</td>
<td>-0,1394</td>
<td>0,0515</td>
<td>-0,0328</td>
</tr>
<tr>
<td>Prazer</td>
<td>0,4648</td>
<td>0,5632</td>
<td>0,0758</td>
<td>-0,2018</td>
</tr>
<tr>
<td>Chato</td>
<td>-0,3789</td>
<td>-0,2619</td>
<td>-0,0988</td>
<td>0,3392</td>
</tr>
<tr>
<td>Nutritivo</td>
<td>0,7978</td>
<td>-0,2176</td>
<td>-0,1422</td>
<td>0,1252</td>
</tr>
</tbody>
</table>

Estimativas finais de comunaldade

<table>
<thead>
<tr>
<th>Satisfaz</th>
<th>Natural</th>
<th>Fibra</th>
<th>Doce</th>
<th>Fácil</th>
<th>Sal</th>
<th>Gratificante</th>
</tr>
</thead>
<tbody>
<tr>
<td>0,5829</td>
<td>0,6176</td>
<td>0,7008</td>
<td>0,6208</td>
<td>0,1765</td>
<td>0,4689</td>
<td>0,6077</td>
</tr>
<tr>
<td>Energia</td>
<td>0,5311</td>
<td>0,4660</td>
<td>0,6722</td>
<td>0,2676</td>
<td>0,3277</td>
<td>0,7641</td>
</tr>
<tr>
<td>Calorias</td>
<td>0,4222</td>
<td>0,4344</td>
<td>0,3730</td>
<td>0,4133</td>
<td>0,7008</td>
<td>0,4588</td>
</tr>
<tr>
<td>Qualidade</td>
<td>0,5596</td>
<td>0,5798</td>
<td>0,3370</td>
<td>0,7197</td>
<td>0,2270</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fator 1</th>
<th>Fator 2</th>
<th>Fator 3</th>
<th>Fator 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaz</td>
<td>0,7057</td>
<td>0,0885</td>
<td>0,1998</td>
</tr>
<tr>
<td>Natural</td>
<td>0,7529</td>
<td>-0,2084</td>
<td>0,0555</td>
</tr>
<tr>
<td>Fibra</td>
<td>0,8214</td>
<td>-0,1156</td>
<td>-0,1200</td>
</tr>
<tr>
<td>Doce</td>
<td>0,0684</td>
<td>0,7020</td>
<td>0,0744</td>
</tr>
<tr>
<td>Fácil</td>
<td>0,2382</td>
<td>0,0634</td>
<td>0,3252</td>
</tr>
<tr>
<td>Sal</td>
<td>-0,0924</td>
<td>0,6855</td>
<td>0,0161</td>
</tr>
<tr>
<td>Gratificante</td>
<td>0,6254</td>
<td>0,0770</td>
<td>0,4240</td>
</tr>
<tr>
<td>Energia</td>
<td>0,6594</td>
<td>0,0786</td>
<td>0,1930</td>
</tr>
<tr>
<td>Divertido</td>
<td>0,1629</td>
<td>0,1765</td>
<td>0,4175</td>
</tr>
<tr>
<td>Crianças</td>
<td>-0,0252</td>
<td>0,0340</td>
<td>0,8502</td>
</tr>
<tr>
<td>Encharcado</td>
<td>0,0330</td>
<td>0,0143</td>
<td>0,0942</td>
</tr>
<tr>
<td>Econômico</td>
<td>0,0686</td>
<td>0,2809</td>
<td>0,4149</td>
</tr>
<tr>
<td>Saúde</td>
<td>0,8287</td>
<td>0,2877</td>
<td>0,0519</td>
</tr>
<tr>
<td>Família</td>
<td>0,0616</td>
<td>0,0554</td>
<td>0,7612</td>
</tr>
<tr>
<td>Calorias</td>
<td>-0,1141</td>
<td>0,6267</td>
<td>-0,0072</td>
</tr>
<tr>
<td>Simples</td>
<td>-0,1466</td>
<td>-0,0622</td>
<td>0,0680</td>
</tr>
<tr>
<td>Crocante</td>
<td>0,0734</td>
<td>0,1458</td>
<td>0,3732</td>
</tr>
<tr>
<td>Regular</td>
<td>0,6132</td>
<td>0,0996</td>
<td>0,0270</td>
</tr>
<tr>
<td>Açúcar</td>
<td>-0,1844</td>
<td>0,8166</td>
<td>-0,0522</td>
</tr>
<tr>
<td>Fruta</td>
<td>0,3764</td>
<td>0,1874</td>
<td>-0,2669</td>
</tr>
<tr>
<td>Processo</td>
<td>-0,2363</td>
<td>0,3737</td>
<td>0,0262</td>
</tr>
<tr>
<td>Qualidade</td>
<td>0,6466</td>
<td>-0,2444</td>
<td>0,2048</td>
</tr>
<tr>
<td>Prazer</td>
<td>0,2444</td>
<td>0,2335</td>
<td>0,3368</td>
</tr>
<tr>
<td>Chato</td>
<td>-0,1646</td>
<td>0,0668</td>
<td>-0,2255</td>
</tr>
<tr>
<td>Nutritivo</td>
<td>0,8315</td>
<td>-0,1764</td>
<td>0,0517</td>
</tr>
</tbody>
</table>

Os atributos que carregam altamente sobre o segundo fator comum são: Doce, Sal, Calorias e Açúcar. Processado também sobrecarrega mais intensamente sobre o fator 2, com uma correlação um pouco abaixo de 0,40. Roberts e Lattin usaram o rótulo “Artificial” para descrever esse fator. Esse rótulo particular é consistente com o fato de os atributos Natural, Saudável e Nutritivo também carregarem negativamente sobre esse fator. Como esse fator não está correlacionado com o primeiro fator, seria inappropriate rotulá-lo como “Não saudável” (que sugere uma correlação forte, se não perfeita, com o fator 1).

Os atributos que carregam mais altamente sobre o terceiro fator comum são Crianças, Econômico e Família. Os atributos Gratificante e Divertido têm também cargas próximas de 0,40 para esse fator (mas possuem cargas mais elevadas sobre outros fatores). Esse padrão geral é consistente com um fator que reflete a “Popularidade” ou “Amplio Apelo” do cereal. Roberts e Lattin usaram o termo “Não Adulto” para refletir o grau em que o cereal não era percebido como estando posicionado exclusivamente no nicho de cereal destinado a adultos.

Por fim, os atributos que carregam mais altamente sobre o quarto fator comum são Divertido, (Encharcado), (Simples), Crocante, Fruta, Prazer e (Chato). Os atributos entre parênteses apresentam cargas negativas; por exemplo, quanto mais encharcado o cereal, menor o valor sobre o fator 4. Roberts e Lattin usaram o termo “Interessante” (talvez tenham sentido que o termo “Empolgante” era exaltado demais para cervejas matinais). De qualquer modo, isso deve fornecer algumas impressões para o processo de interpretar a solução fatorial e nomear os fatores.
O resultado é que reduzimos um problema envolvendo 25 medidas de atributo a um que envolve somente quatro fatores comuns interpretáveis. Fazendo isso, podemos explicar mais da metade da variação nos dados originais.

5.4 QUESTÕES RELATIVAS À APLICAÇÃO DA ANÁLISE FATORIAL

5.4.1 POSSO OBTNER UMA SOLUÇÃO COM FATORES CORRELACIONADOS?

Quando extraímos fatores ou componentes (na análise de componentes principais ou no modelo de fator comum), somos forçados a assumir que haja ortogonalidade (isto é, pressupor por imposição que os fatores comuns não são mutuamente correlacionados) para identificar as soluções. No entanto, uma vez que tenhamos extraído os fatores, não há razão, em teoria (embora haja algumas considerações práticas discutidas a seguir), por que não possamos escolher uma rotação subsequente que permita que os fatores sejam correlacionados. Não limitar que a rotação seja ortogonal pode garantir melhor aproximação à estrutura simples na matriz de cargas transformada e, desta maneira, incrementa-se a capacidade de interpretação da solução. Referimo-nos a essas rotações não ortogonais como rotações oblíquas. A intenção por trás de uma rotação oblíqua é alinhar os eixos do fator tão próximo quanto possível a grupos de variáveis do conjunto original de dados – ou, de outra maneira, facilitar a capacidade de interpretação da configuração resultante –, seja os fatores resultantes correlacionados ou não. Desse modo, a função objetiva para a rotação oblíqua é similar àquela para a rotação ortogonal (por exemplo, minimizar a soma dos produtos cruzados das cargas fatoriais ao quadrado), mas sem a limitação da ortogonalidade.

Em geral, uma transformação oblíqua conduz a várias questões relativas à interpretação da solução fatorial que não são motivo de preocupação quando se usa uma rotação ortogonal. A questão de princípio é que, com uma rotação oblíqua, faz-se necessária uma distinção entre as chamadas cargas de estrutura e cargas de padrão. A Figura 5.7 ilustra a distinção. Suponha que tenhamos algum ponto $p$ no sistema oblíquo que é obtido pela rotação de $A_e$, a solução fatorial ortogonal (ortogonal) da nova matriz $A^* = A_e T^*$, em que $T^*$ representa a matriz de transformação oblíqua. O ângulo $\phi$ reflete a correlação entre os fatores $a^*_1$ e $a^*_2$ no sistema oblíquo.

![Figura 5.7 Diagrama representando a diferença entre cargas de estrutura e cargas de padrão.](image)

Observe que há dois meios pelos quais podemos registrar a posição do ponto $p$ no espaço fatorial oblíquo. Por um lado, é possível registar a projeção (perpendicular) do ponto $p$ nos eixos $a^*_1$ e $a^*_2$, dada por $x_1$ e $x_2$ na figura. No jargão da análise fatorial, há as cargas de estrutura. Por outro lado, podemos registrar as coordenadas do ponto $p$ no sistema de referência oblíquo, dadas por $w_1$ e $w_2$ na figura. Essas são as cargas de padrão. Quando os fatores são ortogonais, a distinção entre cargas de estrutura e cargas de padrão desaparece.

Cargas de estrutura são correlações entre variáveis e fatores (portanto, são limitadas a uma variação entre $-1$ e $+1$). Essas são os valores dados na matriz de cargas transformada $A^* = A_e T^*$. Em geral, as cargas de estrutura não são úteis na interpretação da solução fatorial porque a estrutura simples é disfarçada pelas correlações do fator subjacente. Observe na figura que tanto $x_1$ quanto $x_2$ são relativamente amplos devido à correlação entre os dois fatores. Em contraste, as cargas de padrão são como coeficientes de regressão parcial, no sentido de que levam em consideração a variação explicada por outros fatores. Em outras palavras, elas são coeficientes em uma combinação linear dos fatores obliquos que nos permitem reconstruir a configuração original. Como resultado, as cargas de padrão não são limitadas a variar entre $-1$ e $+1$, mas são geralmente mais úteis na interpretação da solução fatorial. Como apresentado na figura, $w_2$ é maior que $w_1$, o que sugere que o ponto $p$ está mais proximamente alinhado com o fator um. As cargas de padrão são obtidas das cargas de estrutura através da seguinte expressão:

$$A^* = W P$$

onde $P$ é uma matriz padrão de direção dos eixos (apropriadamente normalizados), que capta os ângulos entre os fatores correlacionados no sistema de referência oblíquo. No exemplo bidimensional dado na Figura 5.7, a relação entre cargas de padrão e de estrutura é dada como segue:

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} 1 & \cos \phi \\ \cos \phi & 1 \end{pmatrix}$$

Outra advertência a se ter em mente quando se trabalha com as soluções oblíquas: como os fatores podem estar correlacionados, a comunalidade de uma variável não pode ser computada como a soma das cargas ao quadrado na linha $i$ da matriz de cargas fatoriais. Além disso, a variância explicada por um fator não pode ser computada como a soma de cargas ao quadrado na coluna $k$ da matriz de cargas fatoriais.

Agora, ilustramos a rotação oblíqua voltando ao exemplo ilustrativo no teste psicológico. Examinando a matriz de cargas fatoriais não rodada (apresentada na Tabela 5.5), a interpretação da solução de dois fatores é ainda bastante clara: os três primeiros testes carregam principalmente sobre o primeiro fator (que podemos interpretar como aptidão verbal) e os dois últimos carregam principalmente sobre o segundo fator (que podemos interpretar como aptidão quantitativa). A questão é a seguinte: pode-se obter um retrato mais completo da estrutura simples com a aplicação de alguma rotação oblíqua à solução?

Para responder a essa questão, usaremos uma abordagem à transformação oblíqua conhecida como ajuste de matriz alvo. A matriz alvo consiste em uma representação específica da estrutura simples subjacente à solução do problema. Ela pode ser construída sobre base teórica ou pode representar a estrutura de uma matriz de cargas fatoriais interpretada. Nesse caso particular, nossa matriz alvo é uma matriz simples, dada por

$$G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Nossa meta é encontrar uma matriz de transformação que rodrá a matriz de cargas fatoriais $A_e$ com a congruência mais próxima possível da matriz alvo $G$. Se usarmos $T^*$ para representar a matriz de transformação, nossa meta é escolher $T^*$ para que a transformação resultante $A_e T^*$ seja tão próxima quanto possível da matriz alvo $G$. A transformação que representa o melhor ajuste de mínimos quadrados a $G$ é dada por

$$T^* = (A^*_e A_e)^{-1} A^*_e G$$

O procedimento não é mais elaborado do que regredir cada coluna de $G$ sobre $A_e$ e depois normalizar as colunas resultantes de $T^*$. Observe que as colunas de $T^*$ não serão, em geral, ortogonais. Esse tipo de transformação é, às vezes, conhecido como rotação procrusteanas (veja Hurley e Cattell, 1962).

Os resultados da rotação oblíqua da solução de dois fatores para os dados do teste psicológico são apresentados na Tabela 5.12. As cargas de padrão mostram uma aproximação muito grande à estrutura simples de alvo: as cargas cruzadas no padrão fatorial estão todas próximas de 0,02 a 0,06 em valor absoluto. Portanto, temos uma solução clara na qual o primeiro fator é medido pelos três primeiros testes e o segundo fator é medido pelos dois últimos.
A transformação obliqua resulta em uma correlação entre os dois fatores de 0,25. Essa modesta correlação positiva sugere que os estudantes com alta aptidão verbal (da forma como foi mensurada pelos testes de compreensão de parágrafo, complementação de sentenças e significado das palavras) terão também uma aptidão maior do que a aptidão quantitativa média (e vice-versa). Note que essa correlação entre fatores é refletida nas correlações de estrutura: todas as cargas cruzadas na matriz estão na proximidade de 0,25. Dessa maneira, trocamos um pouco da simplicidade de uma rotação ortogonal por uma solução de fator comprovadamente “mais completa”. Em geral, não é má ideia conduzir tanto a rotação ortogonal quanto a obliqua para determinar qual delas serve melhor aos objetivos da análise.

Uma rotação obliqua dos dados do cereal RTE não oferece nem uma interpretação diferente nem um quadro substancialmente mais completo da estrutura simples. No espaço obliquo, todos os quatro fatores são posicionalmente intercorrelacionados (posivelmente, reflexo de algum efeito halo subjacente), com uma exceção: os fatores 1 e 2 exibem uma pequena correlação negativa. Isso é consistente com a noção de que um cereal percebido como “Saudável” (isto é, alto em fator 1) é mais provavelmente percebido como não “Artificial” (isto é, abaixo da média no fator 2).

5.4.2 Como posso usar os resultados em análises subsequentes?

Com frequência, a análise fatorial não é um fim em si, mas um passo intermediário ao caminho para a análise posterior de dados. Por exemplo, Roberts e Lattin usaram a análise fatorial para identificar um número menor de dimensões subjacentes que poderiam ser usadas na construção de um modelo de consideração para os cereais RTE. Para esse tipo de análise subsequente, é necessário que se conheça a localização de cada observação original no reduzido espaço fatorial. Esses valores são chamados de escores fatais.

Os coeficientes de escore fatorial da solução rodada varimax de quatro fatores para os dados dos cereais RTE são dados na Tabela 5.13. Usamos esses coeficientes para calcular os escores fatais de cada uma das 235 observações no conjunto de dados e depois calculamos os escores médios fatais de cada uma das 12 marcas classificadas. Esses escores médios (apresentados na Tabela 5.14) são esquematizados na Figura 5.8 para economizar espaço, mostramos apenas dois gráficos de pares (Fator 1 versus Fator 2 e Fator 3 versus Fator 4).

Os escores fatais médios são consistentes com nossas noções anteriores acerca das posições relativas desses cereais. Por exemplo, Weetbix e Vitabrits (que têm o formato de biscoito de farelo de trigo) são percebidos de maneira semelhante: saudáveis, naturais e desinteressantes. As três diferentes marcas de musli (Cerola, Komplete e Purina) também são agrupadas no espaço do fator.

5.4.3 Como avaliar a validade da estrutura de fator?

Em todos os procedimentos exploratórios, é preciso avaliar o grau com que nossa análise descritiva capta características da população como um todo e não apenas as características de nossa amostra selecionada. Portanto, precisamos fazer a pergunta: os resultados da análise são generalizáveis? Como a interpretação da estrutura fatorial subjacente é um objetivo importante da análise fatorial, devemos também perguntar: se eu tivesse escolhido uma amostra diferente da mesma população, teria alcançado uma interpretação substancialmente igual?
Figura 5.8 Diagramas de dispersão dos escores fatoriais para cereais RTE.

Pode-se responder a essa questão através do uso de uma amostra de teste. Se o tamanho de nossa amostra de corte transversal for suficientemente grande, podemos simplesmente dividir os dados em dois grupos separados de tamanho aproximadamente igual (utilizando atribuição aleatória para assegurar a comparabilidade). Podemos, então, efetuar uma análise fatorial em cada um e comparar as matrizes de cargas fatoriais por similaridade de interpretação. Uma comparação baseada em inspeção visual casual, no entanto, é raramente satisfatória. Pode haver centenas de números envolvidos, o que torna difícil tal comparação. Ademais, a tarefa é subjetiva e sujeita ao pensamento desejoso (isto é, o pesquisador pode apresentar a tendência de se apegar às semelhanças entre as matrizes, minimizando os contrastes). Uma abordagem mais rigorosa envolve o uso da análise fatorial confirmatória para testar a estrutura fatorial específica identificada na primeira amostra, utilizando-se os dados da amostra de teste para tanto. Se a estrutura fatorial específica não for rejeitada (isto é, o modelo fornece um ajuste razoável aos dados do teste), podemos então concluir que a interpretação se estende para além da primeira amostra de dados. A análise fatorial confirmatória é apresentada e discutida no Capítulo 6.

Outra abordagem envolve avaliar o impacto da variação da amostragem sobre a estrutura fatorial (via uso de uma amostra de teste ou amostragem bootstrap). Vamos supor que usamos o seguinte procedimento para identificar a estrutura subjacente de uma solução com fatores comuns: depois de algum tipo de rotação (vamos dizer variáveis), identificamos a carga mais alta (em valor absoluto) em cada linha da matriz de cargas fatoriais e associamos a variável àquele fator. Então, interpretamos cada fator comum com base nas variáveis associadas a ele (isto é, ou uma elevada carga positiva ou uma elevada carga negativa). Efectivamente, isso levou-nos a substituir a matriz de cargas fatoriais por uma matriz de zeros e 1 indicativa da estrutura simples subjacente à solução fatorial. O resultado é uma divisão na qual cada variável é associada a um e somente um fator comum. Essa "estruturação" da matriz de cargas fatoriais forma a base da nossa interpretação. Usamos $S_0$ para representar essa transformação de estrutura simples da matriz de cargas fatoriais para os dados originais. A matriz $S_0$ para os dados do cereal RTE é apresentada na Tabela 5.15. Note que os atributos têm sido resequenciados para que os atributos que carregem sobre o mesmo fator sejam listados conjuntamente.

Agora, fazemos outra pergunta: com uma amostra diferente da mesma população, teríamos uma interpretação semelhante? Para responder, poderíamos fazer mais uma pergunta: com uma amostra diferente, encontraríamos o mesmo conjunto de variáveis associadas aos fatores comuns, se seguissemos o procedimento descrito no parágrafo anterior? Com uma amostra de teste, podemos fazer isso, conduzindo uma análise fatorial dos dados, extrairíamos fatores (o mesmo número que a solução fatorial nos dados originais) e criando uma matriz de cargas fatoriais de estrutura simples, representada por $S_1$, seguido o procedimento descrito no parágrafo anterior.

Temos agora duas divisões, $S_0$ e $S_1$, cada uma representando uma atribuição de variáveis aos fatores comuns subjacentes. Um modo de compará-los é criar uma tabulação cruzada $c \times c$, em que a célula $(i, j)$ da tabela representa o número de variáveis associadas ao fator comum $i$ na análise dos dados originais e o fator $j$ na análise dos dados do teste. Usando notação matricial, essa tabulação cruzada é dada pelo produto das matrizes $S_0S_1$. Caso a interpretação da análise fatorial original seja válida, ela deve ser estendida para a amostra de teste. Portanto, de se esperar uma forte correspondência entre as partição $S_0$ e $S_1$, com registros elevados na diagonal da tabulação cruzada e registros discretos ou de valor zero fora da diagonal. Há medidas, como a estatística Rand modificada de Hubert e Arabie (1985), para comparar partição, que podem ser usadas para quantificar a força da correspondência entre as duas estruturas fatoriais.

É possível também avaliar a estabilidade de uma interpretação particular de estrutura fatorial através do uso de amostragem bootstrap. Uma amostra bootstrap é simplesmente uma amostra de tamanho $n$ efetuada com a recolocação das $n$ observações originais. Note que algumas observações não serão incluídas na amostra bootstrap, enquanto outras podem ser incluídas duas ou três vezes. A amostra bootstrap nos fornece uma ideia dos efeitos de variação de amostragem que seriam de se esperar ao assumirmos que o conjunto original de dados é representativo da população subjacente. Para cada amostra bootstrap $k$, efetuamos uma análise fatorial e formamos uma matriz $S_k$ de estrutura simples. Depois, avaliamos a estabilidade de nossa interpretação tendo em vista a variação na correspondência entre as matrizes de estrutura $S_0$ e $S_k$ por todas as amostras bootstrap $k$.

Para avaliar a variabilidade através das múltiplas divisões, é mais fácil formar uma tabulação cruzada $p \times p$ de todas as variáveis no modelo, em que a célula $(i, j)$ indica se as variáveis $i$ e $j$...
estão associadas com o mesmo fator. Na notação matricial, essa tabulação cruzada é dada pelo produto de matriz $S_iS_i'$. Podemos efetuar um grande número de amostras bootstrap (digamos, $K$) e então tomar a soma $\sum_{k=1}^{K} S_kS_k'$. A tabulação cruzada resultante contém o número de vezes que cada par de variáveis é identificado com o mesmo fator comum subjacente. Se nossa interpretação for estável (isto é, relativamente não afetada pela variação da amostragem em nossos dados), devemos esperar que as variáveis que carregaram juntas em nossos dados originais também sejam associadas ao mesmo fator comum na maioria ou em todas as nossas amostras bootstrap.

Como ilustração, usamos esse procedimento para avaliar a estabilidade da estrutura simples subjacente à nossa interpretação da solução fatorial para os dados do cereal RTE (apresentados na Tabela 5.15). Efetuamos um total de $K = 25$ amostras bootstrap. A tabulação cruzada dada pela soma dos produtos cruzados da matriz $S_kS_k'$ é mostrada na Tabela 5.16. Os padrões na tabela revelam muitos aspectos da solução fatorial, que não podem ser avaliados olhando-se apenas para a matriz de cargas fatoriais dos dados originais. Em primeiro lugar, observamos que a estrutura subjacente aos fatores 1 e 2 ("Saúde" e "Artificial") é extremamente estável, pelo menos mais do que a estrutura subjacente aos fatores 3 e 4 ("Não Adulto" e "Interessante"). Os 14 atributos associados aos primeiros dois fatores na solução original, todos, menos dois, exibiram exatamente o mesmo padrão de carga da amostra original em todas as 25 amostras bootstrap. Os dois atributos restantes (Gratificante e Processado) são consistentes com a estrutura original em mais de 80% das amostras bootstrap (22 e 21 vezes de 25, respectivamente).

Parece haver mais variabilidade na estrutura subjacente ao terceiro e quarto fatores comuns. Por exemplo, os atributos Crocante e Divertido, que foram associados ao quarto fator ("Interessante") na amostra original, são associados aos atributos Crianças e Família (do terceiro fator comum "Não Adulto") em mais da metade das amostras bootstrap. Além disso, o atributo "Fruta" carrega com as variáveis do quarto fator somente em aproximadamente metade do tempo; no resto do tempo, ele carrega principalmente sobre o primeiro fator (e ocasionalmente sobre o segundo ou o terceiro fator).

Não temos estatística para aplicar aos números da Tabela 5.16 e testar a confiabilidade da estrutura fatorial escolhida. Como muitos dos outros passos da análise fatorial exploratória (por exemplo, escolher o número de fatores subjacentes e interpretar a solução), a avaliação da validade da interpretação fatorial é um julgamento subjetivo (na medida em que ele sequer é praticado!). Mesmo que nossa interpretação substantiva não se modifique, a análise é útil no sentido de revelar a extensão com que a estrutura pode mudar com as variações na amostra.

Pode ser tentador, depois de um exercício de validação como o anterior, “ajustar-se” a solução fatorial. Isso normalmente se dá pelo descarte de variáveis que parecem menos estáveis, em termos de sua posição na solução fatorial, para que assim se chegue a uma solução fatorial mais completamente interpretável. Portanto, no contexto dos dados do cereal RTE, é uma tentativa abandonarmos as variáveis Divertido, Crocante e talvez Fruta para refazer a análise fatorial, esperando uma solução final que evite algumas das potenciais ambiguidades reveladas na solução anterior.

Essa é uma tentativa à qual é melhor resistir, em virtude de pelo menos duas razões. Primeira, a retenção seletiva de variáveis é ainda baseada em padrões de variância na amostra, que consiste em uma combinação de variação atribuível à população subjacente mais a variância associada com essa amostra particular. O ajuste pode melhorar o ajuste da amostra de dentro da solução fatorial, mas põe em risco sua capacidade de generalização para além da amostra. No mínimo, se alguma análise fatorial exploratória for utilizada seletivamente para se reduzir o número de variáveis na análise, a solução fatorial subsequente deve então ser reavaliada com o uso de uma amostra de teste.

A segunda razão é que a estrutura simples nem sempre é uma exigência para se chegar a uma boa solução analítica fatorial. Lembre-se de que o objetivo da análise fatorial exploratória é deixar que os dados sugiram a estrutura subjacente. Uma solução que exiba uma estrutura simples é mais fácil de interpretar, mas nem sempre reflete o que é o mundo.
5.5 RESUMO DA APRENDIZAGEM

- A análise fatorial exploratória, como a análise de componentes principais, é um método que pode ser usado para reduzir-se a dimensionalidade dos dados multivariados e para melhor entendimento dos padrões de associação entre as variáveis. Também é útil para a compreensão dos padrões de correlação nos dados e para identificar os traços ou características subjacentes.

- Na análise fatorial, lançamos mão de um conjunto de pressupostos sobre as fontes de variância em nossos dados; isso é conhecido como modelo de fator comum. O modelo considera que a variância em X_i é atribuída a um número relativamente pequeno de fatores comuns (representados por \( \xi_1, \xi_2, \ldots, \xi_c \)) e um único fator específico (representado por \( \delta_i \)). Assim,

\[ X_i = \lambda_1 \xi_1 + \lambda_2 \xi_2 + \ldots + \lambda_c \xi_c + \delta_i \]

onde \( c \) é o número de fatores comuns no modelo (em geral, muito menor que o número de variáveis em X).

- Os fatores comuns são assim chamados porque também contribuem para a variação em todas as variáveis em X.

- O fator específico \( \delta \) pode ser pensado como uma combinação de erro de variância e variância específica ao fator.

- Para determinar uma solução para o modelo de fator comum, consideramos os seguintes pressupostos:

  - Os fatores comuns não são correlacionados entre si – isto é, corr(\( \xi_i, \xi_j \)) = 0.
  - Os fatores específicos não são correlacionados entre si – isto é, corr(\( \delta_i, \delta_j \)) = 0.
  - Os fatores comuns e os fatores específicos não são correlacionados entre si – isto é, corr(\( \xi_i, \delta_j \)) = 0.

- Supondo que as variáveis X sejam padronizadas (isto é, que \( \text{var}(X_i) = 1 \) para todos os \( i \)), a proporção de variância em \( X_i \) atribuída aos fatores comuns, chamada de comunaldade, é dada por

\[ h^2_i = \Sigma \xi_{ik}^2 = 1 - \theta^2_i \]

onde \( \theta^2_i = \text{var}(\delta_i) \) é a variança do fator específico \( \delta_i \).

- Se as comunaldades \( h^2_i \) são conhecidas, é possível determinar uma solução para o modelo de fator comum, realizando-se uma análise de componentes principais da matriz de correlação \( R \) com comunaldades \( h^2_i \) diagonal. Efetivamente, isso é o mesmo que realizar uma decomposição de valor único na matriz de covariância, após remoção da variância atribuída aos fatores específicos.

- Se as comunaldades não são conhecidas (que é normalmente o caso), é possível aproxima-los usando-se a correlação múltipla quadrática (SMC).

- A solução para o modelo de fator comum é determinada pela orientação do primeiro fator, de modo que ele explique a maior variância possível; do segundo fator, de modo que ele explique a maior variância possível que está sujeita a não ser correlacionada com o primeiro fator comum e assim por diante.

- As correlações entre as variáveis X e os fatores comuns \( \Xi \) são chamadas cargas fatoriais e são representadas por \( \Lambda \).

- As vezes, é desejável reorientar-se a solução fatorial para que a matriz de cargas fatoriais exiba uma estrutura mais simples, fazendo com que os fatores sejam mais fáciles de interpretar.

- Uma solução fatorial exibe estrutura simples quando qualquer variável única é altamente correlacionada com um ou alguns fatores comuns e qualquer fator comum é altamente correlacionado com somente algumas variáveis.

- A solução fatorial é reorientada através de um processo chamado rotação. Dois tipos de rotação são discutidos neste capítulo:

  - A rotação ortogonal preserva a ortogonalidade dos eixos (isto é, os fatores rodados permanecem não correlacionados). Dois métodos amplamente usados para efetuar-a rotação ortogonal são a rotação varimax (que tenta construir a estrutura simples focando-se nas colunas da matriz de cargas fatoriais) e a rotação quartimax (que enfoca as linhas).

  - A rotação obliqua permite uma correlação entre os fatores rodados. Uma abordagem envolve rotar a solução de modo que ela corresponda à matriz alvo que exibe estrutura simples. Alguns cuidados sobre as propriedades das soluções do fator obliquo:

    - Como os fatores comuns são agora correlacionados, as comunaldades não são mais dadas pela soma das cargas fatoriais ao quadrado.

    - Em uma solução obliqua, há uma distinção entre carga de estrutura (isto é, a correlação entre uma variável e um fator) e uma carga de padrão (isto é, a correlação parcial entre uma variável e um fator comum controlando outros fatores comuns).

    - Os escores fatoriais representam as posições das observações no espaço do fator comum. Em razão da indeterminação introduzida pelos fatores específicos, esses escores devem ser estimados. Os coeficientes do escor es fatorial são dados por

\[ B = R^{-1} \Lambda_c \]

em que \( R^{-1} \) é o inverso da matriz de correlação observada e \( \Lambda_c \) é a matriz das cargas fatoriais para a solução fatorial com f fatores comuns. Os escores fatoriais são então calculados como

\[ \Xi = X \cdot B \]

LEITURAS SELECIONADAS

Geral

Rotação

EXERCÍCIOS

5.1 Para fins de ilustração no capítulo, usamos um subconjunto de dados de teste psicológico coletados por Holzinger e Swineford (1939). De fato, Holzinger e Swineford coletaram dados de 145 crianças da 7ª e 8ª séries em um total de 26 diferentes testes psicológicos. Um conjunto maior de nove testes (incluindo os cinco utilizados no capítulo) é apresentado a seguir:
5.3 Sessenta estudantes classificaram 10 marcas de refrigerantes (Coca-Cola, Diet Pepsi, Dr. Pepper, Mt. Dew, Pepsi, Royal Crown, 7Up, Sprite, Diet 7Up, Tab) em relação a quatro atributos (calorias, doçura, sacar a sede e popularidade com os outros) em dois momentos diferentes durante o semestre (setembro e novembro). As variáveis no conjunto de dados são definidas como segue:

<table>
<thead>
<tr>
<th>X₁</th>
<th>X₂</th>
<th>X₃</th>
<th>X₄</th>
<th>X₅</th>
<th>X₆</th>
<th>X₇</th>
<th>X₈</th>
</tr>
</thead>
<tbody>
<tr>
<td>X₁</td>
<td>Calorias (setembro)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₂</td>
<td>Calorias (novembro)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₃</td>
<td>Doçura (setembro)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₄</td>
<td>Doçura (novembro)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₅</td>
<td>Sacia a sede (setembro)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₆</td>
<td>Sacia a sede (novembro)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₇</td>
<td>Popularidade (setembro)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₈</td>
<td>Popularidade (novembro)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A matriz de correlação é dada na Tabela 5.18 (e no arquivo SOFT DRINKS).

Analisar os dados usando a análise fatorial. Quantos fatores há? Como você interpretaria a matriz de correlação?
### Tabela 5.19 Matriz de correlação para os dados SIX_VARIABLES

<table>
<thead>
<tr>
<th></th>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>X5</th>
<th>X6</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X2</td>
<td>0.849</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X3</td>
<td>0.462</td>
<td>0.442</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X4</td>
<td>0.416</td>
<td>0.439</td>
<td>0.909</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X5</td>
<td>0.409</td>
<td>0.360</td>
<td>0.499</td>
<td>0.501</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>X6</td>
<td>0.455</td>
<td>0.334</td>
<td>0.478</td>
<td>0.459</td>
<td>0.862</td>
<td>1.000</td>
</tr>
</tbody>
</table>

### Tabela 5.20 Matriz de correlação para os dados LUXURY_CAR

<table>
<thead>
<tr>
<th></th>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>X5</th>
<th>X6</th>
<th>X7</th>
<th>X8</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X2</td>
<td>0.591</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X3</td>
<td>0.356</td>
<td>0.350</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X4</td>
<td>-0.098</td>
<td>0.072</td>
<td>0.38</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X5</td>
<td>0.573</td>
<td>0.408</td>
<td>0.382</td>
<td>0.062</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X6</td>
<td>0.156</td>
<td>0.232</td>
<td>0.517</td>
<td>0.424</td>
<td>0.303</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X7</td>
<td>0.400</td>
<td>0.414</td>
<td>0.611</td>
<td>0.520</td>
<td>0.401</td>
<td>0.479</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>X8</td>
<td>0.282</td>
<td>0.375</td>
<td>0.512</td>
<td>0.346</td>
<td>0.308</td>
<td>0.463</td>
<td>0.605</td>
<td>1.000</td>
</tr>
<tr>
<td>X9</td>
<td>0.519</td>
<td>0.484</td>
<td>0.467</td>
<td>0.167</td>
<td>0.455</td>
<td>0.311</td>
<td>0.574</td>
<td>0.557</td>
</tr>
</tbody>
</table>

5.7 Em janeiro de 1998, 303 estudantes de MBA foram entrevistados a respeito de suas avaliações e preferências sobre 10 diferentes automóveis. Os automóveis, listados em ordem de apresentação na pesquisa, foram BMW 328i, Ford Explorer, Infiniti J30, Jeep Grand Cherokee, Lexus ES300, Chrysler Town & Country, Mercedes C280, Saab 9000, Porsche Boxster e Volvo V90. Cada estudante classificou todos os 10 carros. Para os fins deste exercício, um carro foi selecionado aleatoriamente por cada um dos estudantes, resultando em um tamanho de amostra de 303 avaliações.

Os estudantes classificaram cada carro em relação a 16 atributos. As primeiras oito questões pediam que os estudantes avaliassem o grau em que cada um das seguintes palavras descrevia um determinado carro (em que 1 = "Extremamente descritiva" e 0 = "Não descreve de maneira alguma"): excitante, confiável, luxuoso, próprio para o ar livre, poderoso, estiloso, confortável e vigoroso. As oito questões seguintes solicitavam que os estudantes classificassem seu nível de concordância com cada uma das seguintes afirmações sobre um determinado carro (em que 1 = "Concordo totalmente" e 0 = "Discreto totalmente"): "Este carro é gostoso de dirigir". "Este carro é seguro". "Este carro tem um grande desempenho". "Este carro é para a família". "Este carro é versátil". "Este carro é esportivo". "Este carro é um carro de alto padrão". "Este carro é prático".

Os dados brutos estão disponíveis no arquivo MBA_CAR. Há 18 variáveis no arquivo, definidos como segue:

- Identidade do estudante (disfarçado)
- Identidade do carro:
  1 = BMW 328i
  2 = Ford Explorer
  3 = Infiniti J30
  4 = Jeep Grand Cherokee
  5 = Lexus ES300
  6 = Chrysler Town & Country
  7 = Mercedes C280
  8 = Saab 9000
  9 = Porsche Boxster
  10 = Volvo V90

### Tabela 5.21 Matriz de correlação para os dados EMOTIONS

<table>
<thead>
<tr>
<th></th>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>X5</th>
<th>X6</th>
<th>X7</th>
<th>X8</th>
<th>X9</th>
<th>X10</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X10</td>
</tr>
<tr>
<td>X2</td>
<td>0.20</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X3</td>
<td>0.08</td>
<td>0.30</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X4</td>
<td>0.13</td>
<td>-0.30</td>
<td>0.37</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X5</td>
<td>0.27</td>
<td>-0.22</td>
<td>0.28</td>
<td>0.80</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X6</td>
<td>0.22</td>
<td>-0.23</td>
<td>0.39</td>
<td>0.84</td>
<td>0.85</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X7</td>
<td>0.22</td>
<td>-0.20</td>
<td>0.45</td>
<td>0.76</td>
<td>0.82</td>
<td>0.92</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X8</td>
<td>0.33</td>
<td>-0.07</td>
<td>0.46</td>
<td>0.67</td>
<td>0.72</td>
<td>0.80</td>
<td>0.83</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X9</td>
<td>0.31</td>
<td>-0.08</td>
<td>0.40</td>
<td>0.55</td>
<td>0.60</td>
<td>0.67</td>
<td>0.77</td>
<td>0.76</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>X10</td>
<td>0.25</td>
<td>-0.21</td>
<td>0.48</td>
<td>0.74</td>
<td>0.68</td>
<td>0.78</td>
<td>0.78</td>
<td>0.7</td>
<td>0.70</td>
<td>1</td>
</tr>
</tbody>
</table>

5.8 Westbrook e Oliver (1991) investigaram os tipos de respostas emocionais para a experiência de consumo. Elles realizaram um estudo de uma amostra de proprietários de carros recém-adquiridos. Um total de 125 questionários foi preenchido com respeito às reações emocionais dos respondentes em relação aos seus automóveis. Os autores usaram a medida DES-II de Izard (1997), que contém 10 subescalas que representam a frequência com a qual os sujeitos experimentam cada uma das 10 emoções relacionadas a seguir:

- X1 Interesse
- X2 Alegria
- X3 Surpresa
- X4 Tristeza
- X5 Raiva
- X6 Desgosto
- X7 Desprezo
- X8 Vergonha
- X9 Culpa
- X10 Medo

A matriz de correlação, reproduzida na Tabela 5.21, está disponível no arquivo EMOTIONS. Em sua opinião, qual é a dimensionalidade do espaço psicológico que contém os vários padrões de resposta emocional? Como você interpretaria as dimensões que você encontrou?