



PRO2514 - Pesquisa Quantitativa em Gestão de Operações

Análise Fatorial

Prof. Dr. Renato de Oliveira Moraes



Análise Fatorial

- A análise fatorial exploratória é uma técnica de redução de dimensionalidade
- Ela agrupa as variáveis colhidas (indicadores, variáveis manifestas) em um conjunto de fatores (constructos, variáveis latentes)
- Assim o número de fatores resultantes é (bem) menor que o número de variáveis originais
- Os fatores gerados se referem aos mesmos objetos da base e dados original (com as variáveis), mas o faz de uma perspectiva de nível mais “alto”
- Há uma “perda” de informação que é compensado pela redução da dimensionalidade do problema



Tamanho da amostra

- Evitar amostras com menos de 50 observações. Preferivelmente acima de 100.
- O número ideal da amostra varia na proporção de 5 a 20 casos para cada variável. Alguns autores afirmam que isso não tem fundamento, é empírico.



Sequência de ações

1. Abrir o arquivo de dados apenas com as colunas desejadas – nome do dataset = AF
2. Estatísticas descritivas: `summary(AF)`
3. Olhando as correlações entre as variáveis: `cor(AF)`



KMO

- O coeficiente KMO (Kaiser-Meyer-Olkin) é uma medida da adequação dos dados da amostra para a análise fatorial.
Caso o valor do KMO for muito baixo, analisar para cada variável o valor do MSA (Measure Sampling Adequacy). Tente eliminar a variável com o menor valor de MAS e verifique se o KMO melhorou.
- Em geral, quando os dados apresentam baixa correlação o problema é identificado aqui, e não no teste de Bartlett.
O teste de Bartlett verifica se a matriz de correlações é diferente da matriz identidade



Valores do KMO

- Acima de 0,8: excelente
- Entre 0,7 e 0,8: bom
- Entre 0,6 e 0,7: regular
- Entre 0,5 e 0,6: fraco
- Abaixo de 0,5: inaceitável



Sequência de ações

1. KMO

- `library(psych)`
- `KMO(AF)`

2. Teste de Bartlett - Bartlett's test for sphericity

- `cortest.bartlett(AF)`



Comunalidades extraídas e Auto Valor

Comunalidade

- É o quanto da variância de uma variável original pode explicar o total da variância de todas as outras variáveis incluídas na análise.
- Ou ainda, que os fatores extraídos em conjunto explicam em % a comunalidade indicada de cada variável.

Auto valor – eigen value

- Representa a quantidade do comportamento conjunto das variáveis originais explicado por um fator



Número de fatores extraídos

- Uma condição de contorno: um estudo, trabalho ou modelo anterior cujos resultados serão comparados
- Eigenvalue (auto vetor) maior do que 1: o poder de explicação do fator extraído é superior ao de uma variável original.
- Scree plot



Rotação ortogonal e não ortogonal

- A rotação dos fatores permite interpretar mais facilmente os fatores extraídos
- Rotação ortogonal – conceitos independentes ou distintos
- Rotação não ortogonal – conceitos relacionados, como, por exemplo, dimensões de um conceito.

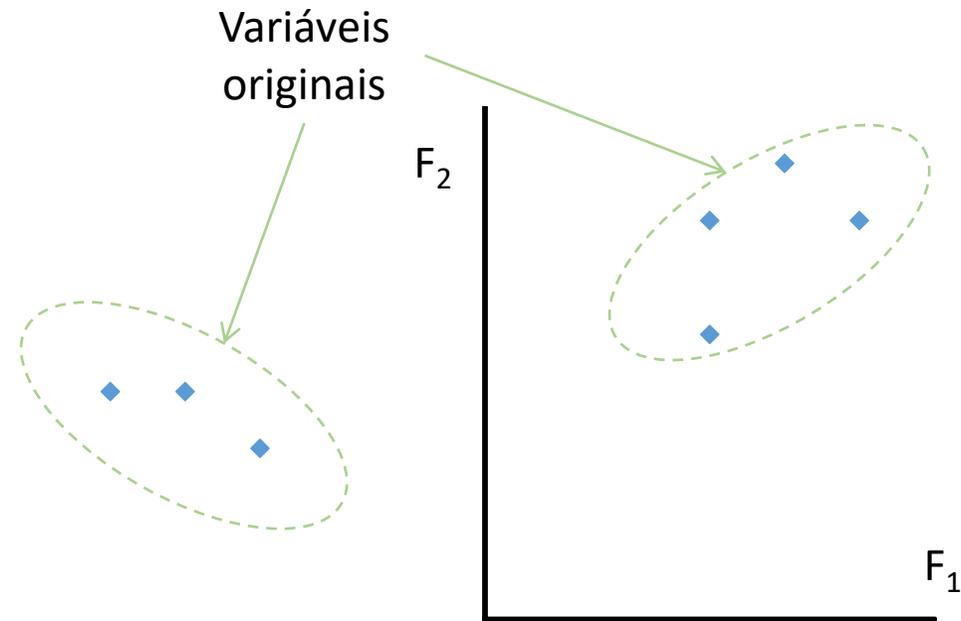


Sequência de ações no RStudio

1. Começando a Análise Fatorial – olhando os eigen values para decidir o número de fatores a extrair
 - `ev <- eigen(cor(mydata)) # get eigenvalues`
 - `ev$values`
2. Desenhar o Scree Plot: `scree(AF, pc=TRUE)`
3. Extração dos fatores: `fit <- factanal(AF, 4, rotation="promax")`
Rotações possíveis:
 - Ortogonais
 - Oblíquas

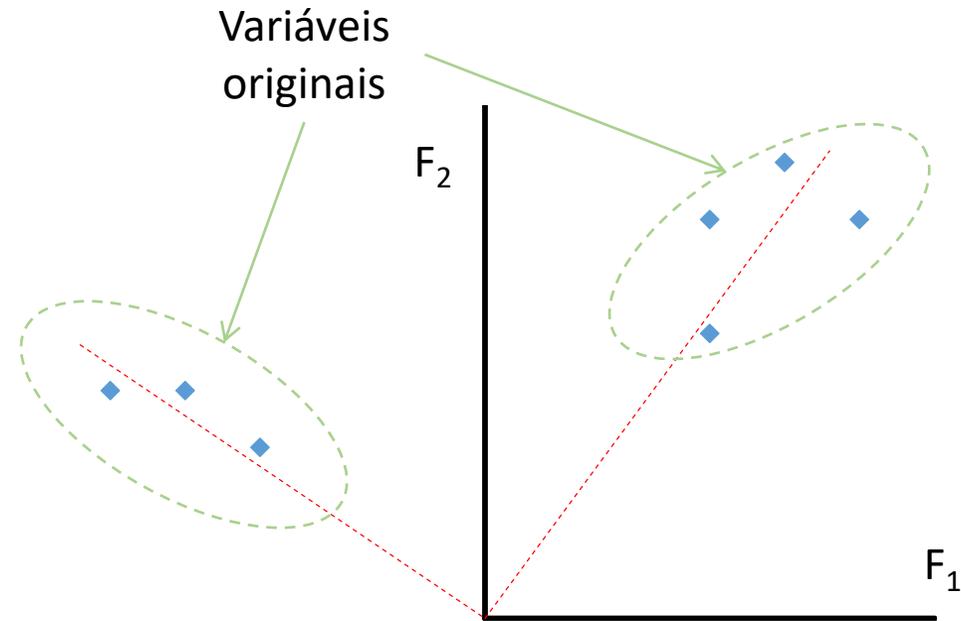


Fatores Extraídos



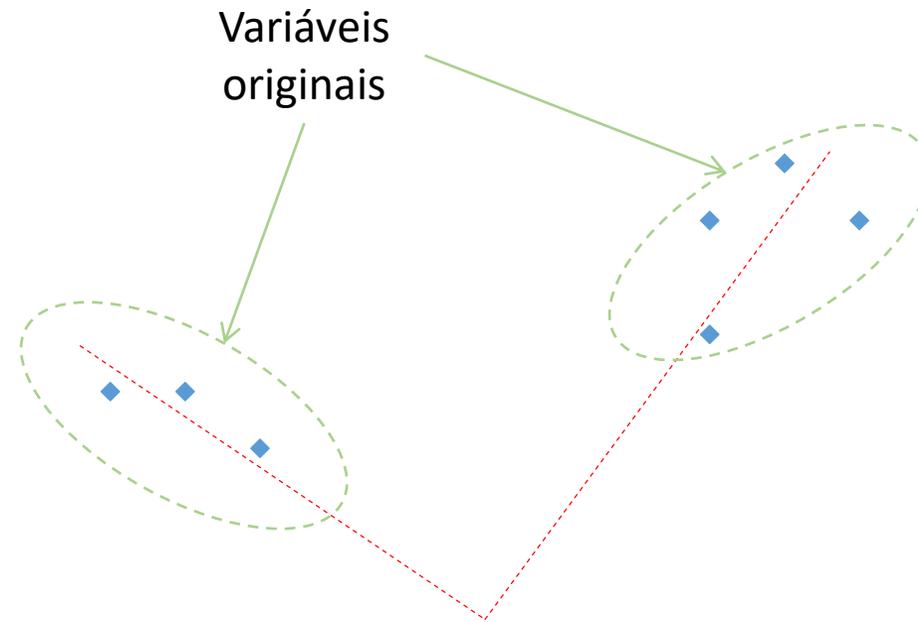


Fatores Extraídos e Fatores Rotacionados



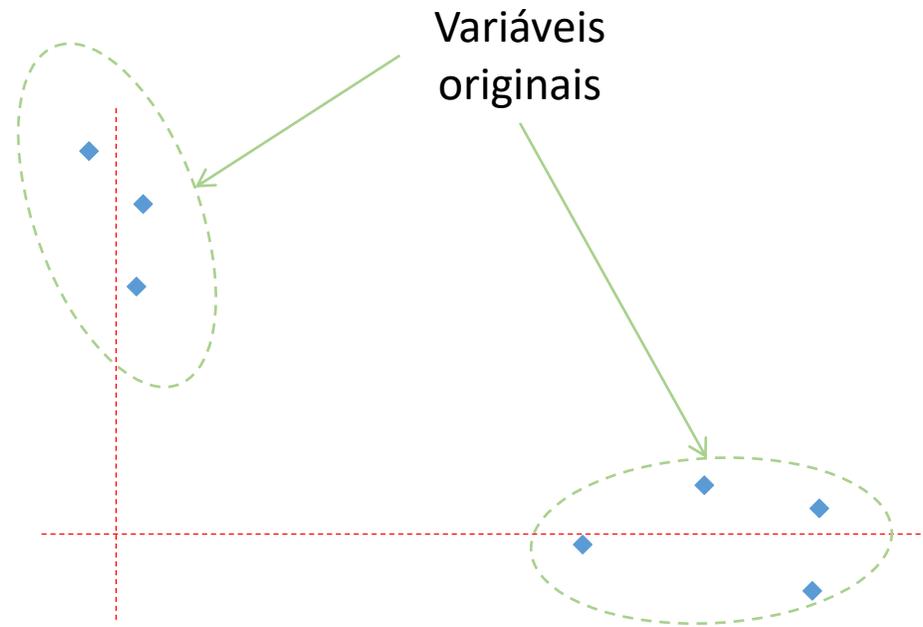


Fatores Rotacionados





Fatores Rotacionados





Cargas fatoriais significativas

Interpretação

- Análise das cargas fatoriais das variáveis
- Considerar na interpretação dos fatores aquelas variáveis cujo valor da carga fatorial seja significativo

Tamanho da amostra	Carga Fatorial
350	0,30
250	0,35
200	0,40
150	0,45
120	0,50
100	0,55
85	0,60
70	0,65
60	0,70
50	0,75



Sequência de ações no RStudio

1. Extração dos fatores: `fit <- factanal(AF, 4, rotation="promax")`
Rotações possíveis:
 - Ortogonais
 - varimax
 - quartimax
 - equamax
 - Oblíquas
 - oblimin
 - Promax:
2. Mostrar resultados: `print(fit, digits=2, cutoff=0.3, sort=TRUE)`



Rotação Ortogonal

- A rotação Varimax é otimizada para reduzir cargas cruzadas e minimizar valores de carga menores, tornando os modelos fatoriais mais claros.
- A rotação Quartimax funciona para reduzir o número de variáveis necessárias para explicar um fator, facilitando a interpretação.
- A opção equamax oferece um compromisso entre varimax e quartimax.



Rotação Oblíqua

- A rotação Promax é popular por sua capacidade de lidar com grandes conjuntos de dados com eficiência. A abordagem também tende a resultar em maiores valores de correlação entre os fatores.
- A abordagem de rotação direta oblímín é um pouco menos eficiente com grandes conjuntos de dados, mas pode produzir uma estrutura fatorial mais simples.



Dar significado aos fatores

- As variáveis colhidas já tinham um significado, mas os fatores ainda não.
- Eles são combinações lineares das variáveis originais. Ao analisar os coeficientes dessas combinações (cargas fatoriais) podemos identificar para cada fator, o peso de cada variável original em seu cálculo
- Os fatores extraídos devem ter significado e utilidade. O pesquisador, conhecendo a teoria subjacente ao tema e aos objetos estudados, deve ser capaz de identificar as implicações e aplicações dos fatores extraídos.



Sequência de ações no RStudio

1. Fazendo um diagrama com os fatores e as variáveis
 - `loads <- fit$loadings`
 - `fa.diagram(loads)`
2. As cargas fatoriais depois da rotação ajudam a dar significado aos fatores: `round(fit$loadings[1:12,], 3)`



Confiabilidade interna – Alpha de Cronbach

- Coeficiente Alfa de Cronbach → análise da consistência interna das variáveis (confiabilidade) indica em que medida as questões foram respondidas de forma coerente.
- Segundo Nunnaly, o ideal é que $\alpha \geq 0,60$.

(Nunnaly, J. Psychometric theory. McGraw-Hill, New York. 1978)



George and Mallery (2003), pag 231

- > 0.9 – Excellent,
- > 0.8 – Good,
- > 0.7 – Acceptable,
- > 0.6 – Questionable,
- > 0.5 – Poor, and
- < 0.5 – Unacceptable

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k s_i^2}{s_{Soma}^2} \right)$$

Onde

α : Alpha de Cronbach

k : número itens na escala

s_i^2 : variância da variável i

s_{Soma}^2 : variância de todas as
variáveis da amostra



Sequência de ações no RStudio

1. Avaliação da confiabilidade

- `f1 <- AF[, c("x1", "x4", "x9")]`
- `alpha(f1, check.keys=TRUE)$total[1]`
- `alpha(f1)`

2. Geração de colunas com os valores dos fatores (escores fatoriais)

- `fit <- factanal(AF, 4, rotation="varimax", scores = "regression")`
- `novos <- as.data.frame.array(fit$scores)`
- `write_xlsx(novos, "Renato/2023/PRO2514 - Pesquisa Quantitativa em Gestão de Operações/saida.xlsx")`