



UNIVERSIDADE FEDERAL DO CEARÁ  
CENTRO DE CIÊNCIAS  
DEPARTAMENTO DE FÍSICA  
CURSO DE LICENCIATURA EM FÍSICA

**FRANCISCO FERNANDO BARROS DOS SANTOS FILHO**

**UMA SIMULAÇÃO COMPARATIVA DO PCA E DO PARAFAC EM  
UM ESTUDO DE CASO DA AVALIAÇÃO DA EFETIVIDADE DO  
ENSINO**

FORTALEZA-CE  
2012



UNIVERSIDADE FEDERAL DO CEARÁ  
CENTRO DE CIÊNCIAS  
DEPARTAMENTO DE FÍSICA  
CURSO DE LICENCIATURA EM FÍSICA

**UMA SIMULAÇÃO COMPARATIVA DO PCA E DO PARAFAC EM  
UM ESTUDO DE CASO DA AVALIAÇÃO DA EFETIVIDADE DO  
ENSINO**

**FRANCISCO FERNANDO BARROS DOS SANTOS FILHO**

Monografia apresentada à Coordenação  
do Curso de Licenciatura em Física da  
Universidade Federal do Ceará, como  
requisito final para a obtenção do grau  
de **Licenciado em Física**.

Orientador: Prof. Ms. Francisco  
Herbert Lima Vasconcelos.

FORTALEZA-CE  
2012

FRANCISCO FERNANDO BARROS DOS SANTOS FILHO

UMA SIMULAÇÃO COMPARATIVA DO PCA E DO PARAFAC EM  
UM ESTUDO DE CASO DA AVALIAÇÃO DA EFETIVIDADE DO  
ENSINO

Monografia apresentada à Coordenação do Curso de Licenciatura em Física da  
Universidade Federal do Ceará, como requisito final para a obtenção do grau de  
**Licenciado em Física**

Aprovada em            de setembro de 2012.

BANCA EXAMINADORA

---

Prof. Ms. Francisco Herbert Lima Vasconcelos (Orientador)  
Universidade Federal do Ceará

---

Prof. Dr. Raimundo Mariano Gomes Castelo Branco (Membro Interno)  
Universidade Federal do Ceará

---

Prof. Ms. Francisca Aparecida Prado Pinto (Membro Externo)  
Secretaria de Educação do Estado do Ceará

Je dédie ce travail a ma copine Joyce Monteiro.

# Agradecimentos

A Deus, o que seria de mim sem a fé que eu tenho nele.

Aos meus pais Albaniza Alves e Francisco Fernando, irmãos, minha namorada Joyce Monteiro, e a toda minha família que, com muito carinho e apoio, não mediram esforços para que eu chegasse até esta etapa de minha vida.

Ao professor Francisco Herbert pela paciência na orientação e incentivo que tornaram possível a conclusão desta monografia.

Ao professor e orientador Albano Oliveira por seu apoio e inspiração no amadurecimento dos meus conhecimentos e conceitos que me levaram a execução e conclusão desta monografia.

Aos amigos e colegas, em especial, Paulo Ricardo e Vanessa Viana, pelo incentivo e pelo apoio constantes.

*“... A avaliação do ensino deveria mostrar-se como orquestra, a fim de apresentar no final do concerto, pelo menos até um certo ponto, uma obra harmônica e com sentido...”*

Bernardete A. Gatti

# Resumo

O presente estudo ensaia uma análise de uma avaliação da eficácia de ensino do curso de licenciatura em Física da Universidade Federal do Ceará, utilizando dados matemáticos que foram gerados usando critérios empíricos educacionais e ponderando-os através do SETE (Avaliação Multidimensional sobre a qualidade de Ensino) e SEEQ (Avaliação Multidimensional sobre a Qualidade Educacional). Serão avaliadas as seguintes disciplinas: Física Fundamental 1, Cálculo Diferencial e Integral 1, Métodos de Física de Fundamental 1 e Estudo Sócio-Político da Educação(ESPE); através das seguintes escalas: aprendizagem, entusiasmo, organização, interação do grupo, afinidade individual, abrangência, exames e trabalhos, tudo isso relacionado a professores e a alunos; já a característica denominada geral estará ligada diretamente as características do curso e da Universidade. As notas atribuídas a cada disciplina são dadas baseadas na Escala Likert, usada geralmente em questionários que possuem respostas psicométricas, e esse tipo de escala é a mais usada em pesquisa de opinião. Os dados foram organizados de forma tridimensional no seguinte arranjo: quatro disciplinas, onze características e duzentos alunos avaliados. Os resultados alcançados foram feitos utilizando dois métodos: PCA (Análise das Principias Componentes), uma análise bidimensional, decomposição PARAFAC (Fatores Paralelos) uma análise multidimensional que visa minimizar o erro, pois para utilização do método do PCA foi necessário ajuste nos dados fazendo a média da dimensão alunos, todavia além de não perder nenhuma dimensão o PARAFAC minimiza o erro através de uma função, comparando os dados estimados com os dados originais. O comportamento dos dados nas duas análises mostrou-se semelhante, por ambas terem a formação de grupo com elementos correlacionados, no entanto não há relação mutua entre os grupos.

**Palavras-chave:** Avaliação Multidimensional; Análise Multivariada; Modelagem Computacional.

# Lista de Figuras

4.1	Desdobramento Matricial . . . . .	31
4.2	Visualização da decomposição PARAFAC. . . . .	33
4.3	Visualização da decomposição Tucker-2. . . . .	37
4.4	Visualização da decomposição Tucker-1. . . . .	38
4.5	Visualização da decomposição PARAFAC. . . . .	38
4.6	Visualização da decomposição PARAFAC é interpretada como um caso especial da decomposição Tucker-3 . . . . .	39
5.1	Representação pictórica dos dados das disciplinas na forma matricial. . . . .	52
5.2	Representação pictórica da redução de dimensão. . . . .	54
6.1	Matriz de coeficientes. . . . .	59
6.2	Matriz de scores. . . . .	59
6.3	Matriz A da decomposição do PARAFAC. . . . .	61
6.4	Matriz B da decomposição do PARAFAC. . . . .	61
6.5	Matriz C da decomposição do PARAFAC. . . . .	62



# Lista de Tabelas

5.1	O conteúdo visto na parte teórica da disciplina . . . . .	50
5.2	Parte experimental é composta pelo conjunto de seis experimentos . .	50
5.3	Conteúdos da disciplina de Cálculo Diferencial Intergral I. . . . .	50

# Acrônimo

Neste trabalho usaremos a seguinte notação: as variáveis escalares serão denotadas por letras minúsculas ( $a, b, \dots, \alpha, \beta, \dots$ ), as matrizes serão correspondidas com letras maiúsculas ( $\mathbf{A}, \mathbf{B}, \dots$ ) e os tensores serão escritas como letras de caligrafia. ( $\mathcal{A}, \mathcal{B}, \dots$ ). O significado dos demais símbolos estão na tabela abaixo:

$\mathbb{C}$	o conjunto dos números complexos;
$\mathbb{C}^I$	conjunto dos números complexos de vetores de dimensão $I$ ;
$\mathbb{C}^{I \times J}$	o conjunto dos números complexos de matrizes;
	com as dimensões $(I \times J)$
$\mathbb{C}^{I_1 \times \dots \times I_N}$	o conjunto do números complexos de tensores;
$a^*$	o conjugado do número complexo $a \in \mathbb{C}$
$ x $	modulus de $a$
$\ x\ _1$	modulo de $a$ $l$ -1
$\ x\ $	modulo de $a$ $l$ -2;
$\mathbf{A}^T$	transposta de $\mathbf{A}$ ;
$\mathbf{A}^H$	transposta Hermitiana de $\mathbf{A}$ ;
$\mathbf{A}^{-1}$	inversa de $\mathbf{A}$ ;
$\mathbf{A}^\dagger$	pseudo-inversa de Moore-Pensore $\mathbf{A}$ ;
$\ A\ _F (\ \mathcal{A}\ _F)$	inversa de $\mathbf{A}$ ;
$\mathbf{1}_N$	todos os elementos do vetor de dimensão $N$ é igual a um;
$\mathbf{I}_N$	matriz identidade de dimensão $N$ ;
$e_n^{(N)}$	$n$ -ésimo vetor canonico em $\mathbb{R}$ , ou seja, um vetor unitário;
	com um elemento igual a 1 na $n$ -ésima posição e 0 em outro lugar;
$E^N = \{e_1^N, \dots, e_N^N\}$	uma base canonica em $\mathbb{R}^N$ ;
$[\mathbf{A}]_{i_1, i_2}$	$(i_1, i_2)$ -ésimo elemento da matriz $\mathbf{A} \in \mathbb{C}^{I_1 \times I_2}$ ;

$[\mathbf{A}]_{i_1.}([\mathbf{A}]_{.i_2})$	$i_1$ -ésima linha ( $i_2$ -ésima coluna) de $\mathbf{A}$ ;
$[\mathcal{A}]_{i_1,i_2,i_3} = a_{i_1,i_2,i_3}$	$(i_1, i_2, i_3)$ -ésimo elemento do tensor $\mathcal{A}$ ;
$\mathbf{A}_{i_1..} \in \mathbb{C}^{I_2 \times I_3}$	$i_1$ -ésima matriz-fatia do primeiro-modo do tensor $\mathcal{A}$ ;
$\mathbf{A}_{.i_2.} \in \mathbb{C}^{I_3 \times I_1}$	$i_2$ -ésima matriz-fatia do segundo-modo do tensor $\mathcal{A}$ ;
$\mathbf{A}_{..i_3} \in \mathbb{C}^{I_1 \times I_2}$	$i_3$ -ésima matriz-fatia do terceiro-modo do tensor $\mathcal{A}$ ;
$a \circ b$	Produto externo entre $a \in \mathbb{C}^{I_1}$ e $a \in \mathbb{C}^{I_2}$
	$a \circ b = \begin{bmatrix} a_1 b_1 & \cdots & a_1 b_{I_2} \\ \vdots & \ddots & \vdots \\ a_{I_1} b_1 & \cdots & a_{I_1} b_{I_2} \end{bmatrix} \in \mathbb{C}^{I_1 \times I_2}$
$\mathbf{A} \otimes \mathbf{B}$	O produto de Kronecker de $\mathbf{A} \in \mathbb{C}^{I \times J}$ com $\mathbf{B} \in \mathbb{C}^{K \times L}$
	$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{1,1}\mathbf{B} & a_{1,2}\mathbf{B} & \cdots & a_{1,J}\mathbf{B} \\ a_{2,1}\mathbf{B} & a_{2,2}\mathbf{B} & \cdots & a_{2,J}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I,1}\mathbf{B} & a_{I,2}\mathbf{B} & \cdots & a_{I,J}\mathbf{B} \end{bmatrix} \in \mathbb{C}^{I \times I_2};$
$\mathbf{A}   \otimes   \mathbf{B}$	O produto de Kronecker "block-wise"
	Para $\mathbf{A} = [\mathbf{A}_{(1)} \dots \mathbf{A}_{(Q)}] \in \mathbb{C}^{I \times J}$ , $\mathbf{B} = [\mathbf{B}_{(1)} \dots \mathbf{B}_{(Q)}] \in \mathbb{C}^{K \times L}$ ;
	$\mathbf{A}   \otimes   \mathbf{B} = [\mathbf{A}^{(1)} \otimes \mathbf{B}^{(1)} \dots \mathbf{A}^{(Q)} \otimes \mathbf{B}^{(Q)}] \in \mathbb{C}^{IK \times R}$ ,
	com $\mathbf{A}^{(q)} \in \mathbb{C}^{I \times J_q}$ , $\mathbf{B}^{(q)} \in \mathbb{C}^{K \times L_q}$ ,
	e $J = \sum_{q=1}^Q J_q$ , $L = \sum_{q=1}^Q L_q$ , $R = \sum_{q=1}^Q J_q L_q$ ;
$\mathbf{A} \diamond \mathbf{B}$	O produto de Khatri-Rao (Kronecker "column-wise")
	para $\mathbf{A} \in \mathbb{C}^{I \times K}$ e $\mathbf{B} \in \mathbb{C}^{J \times K}$
	$\mathbf{A} \diamond \mathbf{B} = [\mathbf{A}_{.1} \otimes \mathbf{B}_{.1}, \dots, \mathbf{A}_{.K} \otimes \mathbf{B}_{.K}] \in \mathbb{C}^{IJ \times K}$ ;

$\text{vec}(\mathbf{A})$	A operação de vetoração Para $\mathbf{A} \in \mathbb{C}^{I \times J} : \text{vec}(\mathbf{A}) = \begin{bmatrix} \mathbf{A}_{.1} \\ \vdots \\ \mathbf{A}_{.J} \end{bmatrix} \in \mathbb{C}^{IJ}$ ;
$\text{diag}(\mathbf{a})$	Matriz diagonal com entradas diagonais dadas pelos elementos de $\mathbf{a}$ ;
bloco diagonal( $\mathbf{A}_1, \dots, \mathbf{A}_N$ )	Forma uma matriz de bloco-diagonal para fora do blocos de matriz $\mathbf{N}$ .
SETE	Avaliação Multidimensional sobre a Qualidade de Ensino
SEEQ	Avaliação Multidimensional sobre a Qualidade Educacional
PCA	Análise das Principais Componentes
PARAFAC	Factores Paralelos
ESPE	Estudo Sócio-Político da Educação

# Sumário

<b>Lista de Figuras</b>	<b>v</b>
<b>Lista de Tabelas</b>	<b>vi</b>
<b>Acrônimo</b>	<b>vi</b>
<hr/>	
<b>1 Introdução</b>	<b>1</b>
1.1 Introdução . . . . .	1
<b>2 Avaliação Multidimensional dos estudantes a cerca da qualidade de ensino</b>	<b>4</b>
2.1 Avaliação em Educação . . . . .	4
2.2 O perfil da análise . . . . .	5
2.3 Generalidade dos SETEs . . . . .	6
2.4 Métodos . . . . .	8
2.5 Implicações . . . . .	10
<b>3 Introdução a Estatística multivariada</b>	<b>14</b>
3.1 PCA - Análise de Componentes Principais . . . . .	14
3.1.1 Definições Principais . . . . .	15
3.2 Análise das Componentes Principais . . . . .	22
3.2.1 Componentes Principais Extraída da Matriz de Covariância . . . . .	22
3.2.2 Critérios para Determinação do Número k de Componentes Principais . . . . .	26
<b>4 Decomposições Tensoriais</b>	<b>28</b>
4.1 Introdução . . . . .	28
4.2 Por Que Utilizar Decomposições Tensoriais? . . . . .	28
4.3 Álgebra tensorial Básica . . . . .	29
4.4 A explanação em decomposição tensorial . . . . .	31

---

<b>5</b>	<b>Procedimentos Metodológicos</b>	<b>49</b>
5.1	Caracterização dos Dados . . . . .	49
5.1.1	Graduações presenciais de Licenciatura Plena . . . . .	49
5.1.2	Conteúdos das Atividades Analisadas . . . . .	50
5.2	Estrutura dos Dados em Três Dimensões . . . . .	51
5.3	Estrutura dos Dados em Duas Dimensões . . . . .	54
5.4	Criação e Tratamento dos Dados . . . . .	54
5.5	Descrição do Processo de Modelagem no MATLAB . . . . .	55
5.5.1	Criando o tensor . . . . .	55
5.5.2	Descrição da Modelagem do PCA . . . . .	55
5.5.3	Descrição da Modelagem no <i>N-way Toolbox</i> . . . . .	56
<b>6</b>	<b>Análise e Discussão dos Resultados</b>	<b>58</b>
6.1	Análise dos Dados . . . . .	58
6.1.1	Análise das Componentes Principais . . . . .	58
6.1.2	Análise Multidimensional . . . . .	60
<b>7</b>	<b>Considerações Finais</b>	<b>63</b>
<b>Apêndice A</b>	<b>Código MATLAB</b>	<b>65</b>
A.0.3	Código Utilizado para Criação das Matrizes . . . . .	65
A.1	Código Utilizado para obter as Componentes do PCA . . . . .	66
<b>Anexo B</b>	<b>Código MATLAB</b>	<b>68</b>
B.0.1	Código Utilizado na Decomposição PARAFAC . . . . .	68
	<b>Referências Bibliográficas</b>	<b>71</b>

# Capítulo 1

## Introdução

### 1.1 Introdução

---

A instrução de classificação do estudante é comumente empregado para a qualidade do ensino na educação. Há um consenso geral de que o aluno pode fornecer informações válidas e opiniões sobre os variados aspectos de ensino. Sendo assim, algumas instituições levam em consideração tais avaliações para que haja promoção de funcionários, no entanto as ferramentas para obter essas informações geralmente estão relacionadas com avaliação seja do aluno, no que diz respeito às notas nas disciplinas ou através de um questionário no qual o aluno atribui notas para determinadas questões estando relacionado com características pré-determinadas. Tais notas geralmente são baseadas na Escala Likert, visto que nesse tipo de avaliação os dados são mais consistentes, possuem mais informações e podem ampliar para outros setores da instituição, enquanto a avaliação referente ao desempenho do aluno acaba sendo restritivo apenas ao aluno e possui um grande empirismo.

A avaliação baseada em questionários tem-se mostrado durante anos uma das mais viáveis. Estudos apontam que é um método seguro, estável e, principalmente, atua em função do instrutor que ensina em um curso diferente do que lhe ensinado, contrário do método baseado no desempenho do aluno, onde há várias variáveis implícitas que geralmente não são levadas em consideração.

Para que houvesse esse tipo análise foi criado um arranjo multidimensional com as dimensões das disciplinas avaliadas, das características do SETE/SEEQ e dos alunos avaliados. Porém como o PCA é uma análise bidimensional, então

fez-se necessário uma redução de dimensão, entretanto no PARAFAC não há perda dimensão, e ao comparar as duas análises, será feita uma, suspostamente, interpretação baseadas nas revisões bibliográficas de SETE/SEEQ. Os resultados obtidos apresentam comportamentos indicando que há um relação entre fatores, mostrando que avaliação com a metodologia SETE pode ser analisada de várias formas, disciplinas que tem características semelhantes encontram-se em um grupo, e as que não tem correlação mostraram-se comportamento contrário.

Os objetivos específicos deste trabalho são:

- i. Realizar uma revisão bibliográfica sobre Avaliação Multidimensional dos estudantes sobre a qualidade de ensino baseando-se nos trabalhos de pesquisadores como: Marsh, H. W., Kulik, J.A., C. Kulik. sobre o SEEQ.
- ii. Realizar uma revisão bibliográfica sobre álgebra linear e multilinear, apresentando suas principais propriedades e características, assim como apresentar algumas ferramentas computacionais importantes no desenvolvimento de trabalhos baseados em decomposições tensoriais.
- iii. Sugerir uma utilização da Avaliação Multidimensional, SETE, para avaliação de desempenho de cursos do ensino superior de forma geral.
- iv. Mostrar o comportamento entre a Análise das Componentes Principais e do PARAFAC.

O trabalho está dividido em sete capítulos:

No **primeiro capítulo** apresentamos uma introdução sobre o tema abordado e como o trabalho está dividido.

No **segundo capítulo** são descritos conceitos teóricos relativos à Avaliação Educacional Multidimensional, através das teorias de Marsh, H. W., Kulik, J.A., C. Kulik. Baseado em algumas exemplificações da utilização da Avaliação.

No **terceiro capítulo** é apresentada uma introdução à álgebra linear, estatística descritiva e PCA descrevendo algumas definições, propriedades, notações e operações básicas que fundamentam essa teoria.

No **quarto capítulo** são mostradas uma pequena introdução de álgebra tensorial e alguma definições matemáticas. São destacadas ainda duas decomposições, as mais conhecidas TUCKER3 e PARAFAC.



No **quinto capítulo** são descritos os procedimentos metodológicos, tais como a caracterização dos dados, as disciplinas analisadas, os conteúdos abordados e a descrição do processo de modelagem para a análise estatística descritiva, bem como a descrição da modelagem no *N-way Toolbox* e do comando *princomp* serão abordadas.

No **sexto capítulo** São apresentados os resultados obtidos através do PCA (Análise das Principais Componentes) para as disciplinas e para as características do SEEQ, no PARAFAC além de mostrar as disciplinas e as características será mostrada o comportamento do alunos em relação ao questionário e às disciplinas.

Por fim, no **sétimo capítulo** serão apresentadas algumas considerações finais obtidas através do estudo teórico.

# Capítulo 2

## Avaliação Multidimensional dos estudantes a cerca da qualidade de ensino

### 2.1 Avaliação em Educação

---

Este texto apresenta algumas questões básicas da avaliação educacional, quais sejam: suas finalidades e suas concepções e a necessária entre ambas. As duas questões nos parecem essenciais, pelo menos por duas razões: a primeira diz respeito ao fato de a avaliação estar profundamente implicada com a função social da educação, ou seja, com o projeto de sociedade que se quer; a segunda relaciona-se, especificamente, ao próprio contexto no qual as práticas avaliativas se desenvolvem, pois compreendem tanto os sujeitos, professores(as) e alunos(as), quanto às ações educativas entendidas aqui como os planejamentos, atividades, mediações, tempos e espaços de aprendizagem.

A avaliação é considerada um elemento de nossos fazeres e ações, pois sem ela não tomamos decisões acerca das questões que compõem nosso cotidiano. Quando a avaliação é realizada de maneira despreziosa, por exemplo, corre o risco de fazermos julgamentos acerca daquilo que está em análise e tomarmos decisões nem sempre adequadas. Ou até mesmo o contrário, podemos ser tão exigentes e fazer um julgamento tão rigoroso que nossas decisões, balizadas por tal rigor, podem comprometer nossas futuras ações acerca daquilo que foi o foco de avaliação.

Enfim, o que é demarcado inicialmente, é que julgamento e decisão são aspectos constitutivos do ato de avaliar, mesmo que esse ato seja desprezencioso. No âmbito educacional e mais especificamente o escolar, a simples constatação de que a avaliação envolve esses dois aspectos faz com que a reflexão acerca do seu papel educativo, social e pedagógico seja mais do que desejada, também seja necessária. Pois, se avaliar implica uma ação que envolve julgamento com vistas a uma tomada de decisão, essa ação deve estar balizada por princípios, sobretudo éticos, uma vez que as decisões decorrentes da avaliação da aprendizagem implicam encaminhamentos na vida escolar dos estudantes.

## 2.2 O perfil da análise

---

A literatura em avaliação multidimensional dos estudantes sobre a qualidade de ensino SETE consiste em milhares de estudos datados de 1920 no qual Marsh & Dunkin (1991) concluíram que o SETE: (a) é multidimensional, (b) é seguro e estável, (c) é principalmente uma função do instrutor que ensina em um curso diferente do que é ensinado, (d) é relativamente válido comparado à variedade de indicadores de qualidade de ensino, (e) não é relativamente afetado pelas variedades de variáveis hipotéticas como a influência do potencial nas avaliações e (f) mostrou ser muito útil para faculdades como avaliação sobre o ensino, feitas aos estudantes no qual influência nas decisões dos administradores do curso e pesquisadores. Baseado na sua revisão, Marsh (1991) argumenta que o SETE provavelmente é o método mais estudado de todos, em relação à avaliação pessoal e um dos melhores, em termos apoiado por uma pesquisa empírica. Os SETE's representam o ensino de forma multidimensional, o professor pode ter característica não correlacionadas, por exemplo, o professor pode ser organizado em seu modo de lecionar, porém não há entusiasmo.

Essa alegação é apoiada por senso comum e é considerado como o corpo de pesquisa empírica. A literatura SETE contém muitos exemplos de instrumentos bem construídos com a definição clara dos fatores de estruturas que são providos de medições de distintas escalas dos SETEs. Murray (1990) notou que a abordagem sistemática usada no desenvolvimento deste instrumento e a similaridade nos fatores identificados por cada apoio deles servem para construir uma validade.

As avaliações dos estudantes sobre a qualidade educacional (SEEQ) são instrumentos que apresenta uma medição amplamente representada por um conjunto

de escalas e têm o fator de maior apoio analítico destes instrumentos. Marsh e Hocevar (1991) analisaram as respostas dos fatores do SEEQ de um grupo no total de 24.158, média das avaliações do conjunto de classe e para cada um dos 21 subgrupos selecionados através de diferentes termos: rank do instrutor, nível do curso (estudante universitário e estudantes que concluíram) e disciplina acadêmica, porém o conjunto dos nove fatores que o SEEQ é designado para medir foi identificado, ao todo, 22 fatores de análise; e o fator de escores está baseado no grupo como um todo, onde está fortemente correlacionado com o fator de escores da análise feita separadamente com os 21 subgrupos (média,  $r > 0.99$ ). Os resultados demonstraram que através do SEEQ obteve-se um conjunto amplamente representativo dos fatores de avaliação e apoiou a generalização da estrutura fatorial do SEEQ.

Uma extensão lógica da multidimensionalidade dos SEEQs é que determinado instrutor tem um perfil diferente da pontuação do SETE e isso se generaliza com o passar do tempo e em diferentes cursos. Tendo assim, segundo o artigo *Ratings of Teaching Effectiveness by Students from Different Academic Settings and Their Relation to Student/Course/Instructor Characteristic* (1983) como ponto de partida para cada pesquisa, faz-se necessário o instrutor tem o perfil dos SETEs distinguível. Utilizando-o como *feedback* para assim em outras relações e em outros critérios.

## 2.3 Generalidade dos SETEs

---

Estudos de generalidade dos SETEs têm tipicamente considerado avaliações globais, total de escores, ou análise separada de dimensão específica dos SETEs. Pesquisadores examinaram correlações entre a pontuação do mesmo instrutor em diferentes disciplinas do mesmo curso, sendo o mesmo instrutor em diferentes cursos e diferentes instrutores ensinando o mesmo curso em uma tentativa de desembaraçar a relativa influência do curso e do instrutor. A pesquisa examinou a generalidade dos SETEs e a relativa importância do instrutor que ensina o curso e o curso que está sendo ensinado. "Para o conjunto de instrutores e avaliação do curso, a correlação entre avaliação de diferentes instrutores ensinando o mesmo curso (uma estimativa da qualidade do curso) onde -0.05 e -0.01, respectivamente, enquanto que a correlação entre avaliação para o mesmo instrutor em duas diferentes classes (0.61 e 0.59) e em dois diferentes oferecimentos do mesmo curso (0.72 e 0.71) onde é muito maior". Estes resultados demonstraram onde são generalizável e é principalmente em função

do instrutor ao invés do curso" (MARSH, 1981).

A recente extensão da pesquisa, *The CIPP Models-viewpoints on Educational Evaluation* (1993), a estabilidade dos SETEs foi avaliado durante 13 semestres. Embora conceitualmente semelhante em pesquisas anteriores destinada a distinguir entre os efeitos do curso e os efeitos dos professores, este estudo é focado no quadro da interação a partir do social e da personalidade psicológica.

De acordo com esse *framework*, o desempenho é uma função pessoal (o instrutor) e a situação (o curso que está sendo ensinado). A fim de testar esta hipótese, Hanges et al. (STUFFLEBEAM, 1993) correlacionou: (a) avaliação do mesmo instrutor (média entre todos diferentes cursos) em 13 semestres diferentes - correlações de "pessoas": (b) avaliação do mesmo curso (média entre todos diferentes instrutores) em 13 semestres diferentes - correlações de situações: e (c) avaliação do mesmo instrutor no mesmo curso - correlações "situação pessoal".

Os resultados eram em grande parte consistente com esses informados por Marsh (1981) em que a correlação "pessoas em situações" era modestamente maior que a correlação "pessoa", considerando que a correlação "situações" era substancialmente pequena. Os autores concluíram que "a fraca estabilidade dos coeficientes para a situação sugere que: pelo menos para conduta do professor na sala de aula, o corpo docente atribui ao invés de atributos a disciplinas ministradas a estabilidade no rendimento em avaliações" (STUDENTS'..., 1984). Porém, calculando a média por todas as avaliações do mesmo curso em cada semestre, obteve-se um efeito contrario da situação. Na verdade confundiu com a avaliação do mesmo curso ensinado por diferentes instrutores (uma estimativa do efeito "pessoa e situação"). Conseqüentemente, embora a estimativa deles da situação fosse muito pequena, era positivamente parcial e uma estimativa imparcial teria sido até menor. Discutindo limitações existentes na pesquisa e direções para estudo futuro, Marsh (STUDENTS'..., 1987) notou que a generalidade da média da avaliação de separadas escalas do SETE foram examinadas, mas não a generalidade do perfil do SETE. Considerando que Hanges et al (1981) avaliaram separadamente diferentes escalas do SETE.

Nenhuma tentativa foi feita, por exemplo, para determinar se o instrutor é entusiástico, e, no entanto falta organização para um curso receber um padrão consistente da avaliação em relação a outros cursos. Marsh (STUDENTS'..., 1987) notou que muitas pesquisas eram necessárias para determinar a generalidade dos

perfis de escores como também a generalidade de cada escore dentro do perfil. Ele discutiu que a existência de diferenças individuais seguras nos perfis da dimensão SETE teria importância teórica e implicações práticas. Se cada instrutor tem um perfil distinguível que é distinto a partir do perfil típico de outros instrutores, então as pesquisas têm que examinar a relação inteira de perfil de escores de diferentes critérios de validade (por exemplo, aprendizagem do estudante) como também a relação de cada escore dentro do perfil. Não conhecemos nenhuma pesquisa e isso fez com que procurássemos este assunto.

Segundo Murray (1990) importante distinguir entre o seu nível (por exemplo, se a avaliação SETE por todas as escalas possuem alta consistência ou baixa para um instrutor particular) e a sua forma (por exemplo, se cada instrutor tem um distinguível perfil de escores).

## 2.4 Métodos

---

Durante os 13 anos, no período de 1976 - 1988, quase um milhão de formas de SEEQ foram administradas em aproximadamente cinquenta mil cursos de uma grande universidade particular no oeste dos Estados Unidos. Embora o uso da SEEQ seja de forma voluntária, a universidade requereu que cada unidade acadêmica coletasse algumas formas dos SETEs para todos os cursos. O quadro de funcionários não o considerou como avaliação para promoção. Muitas das unidades acadêmicas que usaram o SEEQ solicitaram que todo o seu pessoal fosse avaliado em todos os cursos.

As formas de avaliação são tipicamente distribuídas aos funcionários pouco antes do final de cada período letivo, administrados e recolhidos por um estudante ou por um membro do suporte administrativo de acordo com as ordens impressas e levadas a um escritório central onde elas são processadas. Este programa, o instrumento SEEQ em que se baseia, e pesquisas que levaram ao seu desenvolvimento são descritos por Marsh (1992). Ambas medidas repetidas e abordagens multivariadas fazem um perfil do "endereço" da análise das três questões:

- i. Nível. Há uma diferença sistemática de instrutor nos escores total do SETE isto é, a média de todos os nove escores do SEEQ. Na tradicional análise de medidas repetidas, o total de escores que o mesmo instrutor recebeu foram consistentes por diferentes cursos oferecidos e variam sistematicamente a partir do total de

escores por outros instrutores. Este tipo de questão é tipicamente procurado em estudos prévios da generalidade das respostas do SETE. Embora este não seja o foco primário da presente investigação, é uma questão muito importante.

- ii. Paralelismo. São os perfis da escala do SEEQ de diferentes instrutores paralelos ou fazer com que os perfis variem sistematicamente para diferentes instrutores a serem destacados. Na análise tradicional dos dados repetidos que se refere à diferença nas escalas do SEEQ (a repetida medidas de variáveis) faz com que interaja com a variável do instrutor. O efeito testa se há uma diferença individual sistemática nos perfis dos escores do SEEQ. Consequentemente, a significância e tamanho deste efeito são de maior importância para o resultados deste estudo.
- iii. Igualdade. Este item questiona se há diferenças sistemáticas nas nove contagens de SEEQ sendo calculadas a média entre todos os instrutores. Esta questão, para esse trabalho não interessa muito, como notado por Tabanick e Fidell (1990) a questão é tipicamente irrelevante a menos que a suposição de paralelismo seja aceita. Como sugerido por Stevens(1989) e por Tabachnick e Fidell, todos os escores do SEEQ são padronizados (média = 0, SD =1) para eliminar este efeito. Porque os escores do SEEQ foram baseados nos escores dos fatores, no qual já foram padronizados. Mesmo se eles não fossem padronizados, no entanto, as comparações, com dados brutos, das pontuações médias para diferentes escalas não se justificaria. Assim, por exemplo, a comparação qualquer de duas escalas para um dado professor seria refletido através de uma combinação desconhecida das características do ensino medido por cada escala e em particular, a formulação dos itens reflete em cada escala.

Primeiro, padronizando todas as escalas, por uma amostra normativa larga, e a comparação reflete como um determinado professor compara com a amostra total em cada uma das duas escalas.

Para propósito da presente investigação, análises foram conduzidas usando em ambos os métodos medidas repetidas e abordagem MANOVA. E entre os fatores estão, avaliados a influência do instrutor, o nível do curso e a interação entre eles nos escores totais do SEEQ (isto é, o total das escalas do SEEQ). De acordo com (SCRIVEN, 1993) um grande efeito do instrutor, por exemplo, indicaria que o total de escores do SEEQ é largamente influenciado pelo instrutor no qual está sendo

avaliado, com foi encontrado em pesquisas anteriores (veja a discussão facilmente no perfil "nível").

O principal efeito dentro do fator foram necessariamente os zeros devido a todas as escalas do SEEQ onde foram padronizadas para terem o mesmo significado e o desvio padrão (veja facilmente do perfil "igualdade"). A interação "dentro" e entre os fatores avaliados, se o efeito dentro os fatores variaram, dependendo da escala do SEEQ. Um grande instrutor por escalas de interação, por exemplo, indicaria que há uma diferença sistemática entre instrutores nos perfis da escala do SEEQ (veja facilmente a discussão no perfil "paralelismo").

Devido ao conhecido problema de avaliação do significado estatística de análise univariada através das medidas repetidas, o Greenhouse-Geisser Epsilon, o Huynh-Feldt, e o no máximo conservador Lower-Bound Epsilon (1983).

## 2.5 Implicações

---

Foi constata no estudo que há uma grande e sistemática diferença nos perfis do SEEQ obtidas por diferentes instrutores. O perfil SEEQ para cada instrutor generaliza os conjuntos de avaliações obtidas pelo mesmo instrutor durante o período de 13 anos durante os cursos de pós-graduação e graduação. Embora estes resultados sejam superficialmente similares, eles são bastante separados e estatisticamente independentes.

O grande efeito dos instrutores no total de escores, a generalidade dos SETEs em múltiplos cursos oferecidos, não diz nada sobre a dimensionalidade das avaliações. O largo efeito dos instrutores nas análises dos perfis, entretanto, provê um forte apoio para a dimensionalidade dos SETEs.

Em resumo, os resultados sugerem que os instrutores apareçam tendo distintos perfis de pesos e a debilidade que são altamente generalizáveis e que os estudantes são aparentemente capazes de discriminar os pesos e as debilidades dos instrutores, pelo menos quando se agregam avaliações segundo muitos estudantes.

A presente investigação é aparentemente o primeiro teste de generalidade dos perfis SETE e neste assunto provê importantes novas evidências em manter a multidimensionalidade dos SETEs. A apresentação dos resultados enfatiza os efeitos separados entre e dentro das seções de análises. Também, o tamanho do efeito para os escores total e para perfis baseados em cada porção da variância em duas seções



diferentes de análises. Portanto, quando o número de escores em um perfil é grande na presente investigação a comparação dos dois é um pouco não confiável. A total soma dos quadrados para dentro dos efeitos é substancialmente maior do que para entre os efeitos. Logo, a consistência nos perfis dos instrutores explica muito mais variância do que a consistência do total do escores dos instrutores - até mais do que é incluído por uma comparação dos respectivos tamanho dos efeitos.

Como é caso com nenhum efeito de interação, estes resultados indicaram que a diferenças entre os instrutores variam substancialmente dependendo da escala específica da escala SEEQ e isso que está sendo considerado, logo devem ser interpretadas diferentes dos escores total e cautelosamente. Esta conclusão é consistente com uma ênfase na multidimensionalidade dos SETEs e a importância nos perfis do mesmo.

Porque há diferenças individuais seguras em perfis do SETE que são independentes dos seus níveis, nossos resultados implicam que os pesquisadores também deveriam considerar como diferentes perfis relacionados com critério externos da qualidade de ensino. Por exemplo, pode ser que aprendizagem do estudante seja maximizada quando em ambos os entusiasmo e organização são altos, considerando que há poucos sendo alto ou outros são insuficientes. Similarmente, a demonstração das sistemáticas diferenças individuais seguras em perfis SETE mantém o uso dos perfis em avaliação formativa e resume as avaliações e tem importantes aplicações para o estudo do ensino e estilo do ensino.

Há(MARSH, 1981) aparentemente uma regra bem definida que as avaliações para uma construção dos conjuntos de múltiplas dimensões dos SETEs assim com os fatores do SEEQ que são utilizáveis mais as avaliações globais ou o total de escores das avaliações para propósitos de respostas consistentes. As respostas, a partir das avaliações globais, formam um indicador da efetividade a partir da perspectiva dos estudantes, mas fornecem um pequeno ou nenhum valor dos diagnósticos sobre quais as áreas específicas que estão em falta de melhoria. As respostas tipicamente consistem da média das avaliações para cada item ou cada componente da efetividade pedagógica, acrescentar, talvez, através de informação sobre a distribuição das avaliações por estudantes individuais. Estas avaliações, porém, proporcionam uma informação limitada. Por exemplo, uma avaliação de 3.5 em uma escala de 5 para organização não responde se minha organização é boa ou se necessita melhorar. Além disso, uma pontuação de 3.5 em organização e uma 3.75 entusiasmo, também não

falam em qual área eu sou mais forte, sou melhor, devido à escala não mostrar significativamente, a característica do item escolhido para representar em cada escala.

A utilidade desta informação é aumentada pela comparação com um conjunto construído adequadamente de avaliação normativa. Assim, por exemplo, se os significados das avaliações nos cursos tal como esse que eu ensinei é 4.0, então minha avaliação de 3.5 pode sugerir que eu necessite melhorar minha organização. Particularmente quando comparação normativa é apresentada, há uma ênfase com as avaliações comparadas com os de outros professores.

O uso apropriado de SETEs para avaliações somativas - particularmente se eles são usados como uma decisão pessoal - é muito controverso da forma como usam "uma resposta formativa". Entretanto, há um debate aquecido sobre se é o método mais apropriado para usar avaliações específicas, avaliações multidimensionais como esses, resultando a partir do SEEQ, ou uma combinação dos dois. A um extremo Abramim e d'Apolonia (1983), e isso foi discutido que apenas a avaliação global poderia ser usado, embora também Abramim tenha notado que a avaliação global tem valores limitados para propósito formativo.

No outro extremo, Frey argumentou que apenas dimensões específicas poderiam ser usadas, e ele especificamente excluiu avaliações globais do instrumento de sua avaliação. (MARSH; DUNKIN., ) recomendou um fundamento superficial no qual ele discute a utilidade de ambas as avaliações globais e da avaliação de dimensões específicas. Avaliação Global reflete em uma média ponderada em que estudantes determinam a importância das diferentes componentes. (MARSH; DUNKIN., ) justificou "porque esta é a valiosa informação a avaliação global SEEQ contém e eu apoio o uso deste". Porém, é consistente com as perspectivas, (MURRAY H. G., 1990) também argumentada "que há muitos critérios de qualidade de ensino, que cada critério será relatado de forma diferenciada em diferentes dimensões do SETE, e que qualquer critério pode ser deduzido com mais precisão com um conjunto equilibrado de dimensões específicas do que com itens da avaliação global".

O uso da média ponderada de dimensões específicas pode representar um compromisso que é consistente com recomendações por Abrami, por Frey, e por Marsh. (PEREYRA; SCHERER, 1973) Diferentes pesos apropriados para um propósito particular poderia ser determinado empiricamente baseado em pesquisas ou reflete um consenso de opiniões de expert baseado na teoria e na prática. Alternativamente,

um perfil de importância relativa de cada dimensão poderia ser estabilizado pelo instrutor que está sendo avaliado. Por exemplo, à apoio da minha aplicação para promoção, poderiam me pedir que defendesse que considero de relativa importância das dimensões do SEEQ em relação a minha própria filosofia de ensino e de circunstância específicas em classes que eu ensino. (STUFFLEBEAM, 1993) A união entre o perfil de minha importância avaliação e meu atual perfil dos SETEs poderia ser um indicador da minha qualidade de ensino - junto com, por ventura, avaliação global ou médias ponderadas de dimensões específicas e outros indicadores não baseado em SETEs. (MARSH; DUNKIN, 1991) Esta proposta tem um benefício somado de provendo com um explícito papel na interpretação dos SETES que são usados em minha promoção e assim podem também aumentar os valores formativos das avaliações.

# Capítulo 3

## Introdução a Estatística multivariada

### 3.1 PCA - Análise de Componentes Principais

---

A técnica denominada de análise de componentes principais, popularmente chamada de PCA, tendo, em 1901, como precursor Karl Pearson foi fundamentada no famoso artigo de Hotelling em 1933 e sua finalidade fundamental é explicar a estrutura de variância e covariância de um vetor aleatório, composto por  $p$ -variáveis aleatórias, construído através de combinações lineares das variáveis originais, essas combinações lineares são chamadas de componentes principais mas não são correlacionadas entre si. Para descrever algo precisa-se de  $p$ -variáveis para obter  $p$  componente e assim descrever algo e interpretá-lo, com o uso do PCA tem-se uma redução do número de variáveis a serem avaliadas e interpretação das combinações lineares construídas, logo a informação contida nas  $p$ -variáveis originais serão substituídas pela informação contida em  $k$ , onde ( $k < p$ ), componentes principais não correlacionadas. A qualidade da aproximação depende do número de componentes mantidas no sistema e pode ser medida através da avaliação da proporção de variância total explicada por essas. Um dos indícios em que o uso do PCA é necessário, da-se quando a distribuição de probalidades do vetor aleatório em estudo é normal  $p$ -variada. As componentes principais, além de não serem correlacionadas, são independentes e têm distribuição normal, porém a suposição de normalidade não é requisito necessário para que a técnica de análise de componentes principais possa ser utilizada, para obter as componentes principais deve envolver a decomposição da matriz de covariância do vetor aleatório de interesse.

### 3.1.1 Definições Principais

#### Estatística Descritiva

A ideia desta seção é introduzir algumas notações e conceitos da estatística descritiva sobre vetores aleatórios, matrizes de covariância e de correlação e os estimadores mais comuns de parâmetros de distribuições multivariadas. Serão apresentadas algumas definições usados no cálculo de dados multivariados:

**Definição 1** (Vetor Aleatório). *Seja  $X$  um vetor contendo  $p$  componentes, onde cada componente é uma variável aleatória, isto é,  $X_i$  é uma variável aleatória,  $\forall i = 1, 2, \dots, p$  e é denotado por:*

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$$

*O vetor transposto do vetor aleatório  $X$  é denotado por  $X' = [X_1 X_2 X_3 \dots X_p]$ .*

**Definição 2** (Vetor de médias). *Seja  $X$  um vetor aleatório. O vetor  $\mu = E(X)$  é chamado de vetor de médias do vetor  $X = [X_1 X_2 X_3 \dots X_p]'$ , sendo*

$$\mu = E(X) = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$$

*onde  $\mu_i = E(X_i)$  denota a média, ou esperança, da variável aleatória  $X_i$ , onde  $i = 1, 2, \dots, p$ . A média  $\mu_i$  é uma das medidas mais utilizadas para sintetizar a informação de tendência central da distribuição de valores da variável  $X_i$*

**Definição 3** (Variância). *A variância de uma variável aleatória indica quão longe em geral os seus valores se encontram do valor esperado; Variância sempre será diferente de zero; Não depende da média;*

$$\sigma_i^2 = \sigma_{ii} = Var(X_i) = \frac{1}{N} \sum_{i=1}^N (y_i - \mu)^2, \quad \text{onde } i = 1, 2, 3 \dots N \quad (3.1)$$

$$S_i^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad \text{onde } i = 1, 2, 3 \dots n \quad (3.2)$$

*Equação (1) variância populacional, Equação (2) Variância amostral e  $\mu$  é a média populacional.*

**Definição 4** (Desvio-Padrão). *É uma das medidas mais antigas da estatística usado para medir a dispersão estatística, definido com a raiz quadrada da variância.*

$$S_i = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}, \quad \text{onde } i = 1, 2, 3 \dots n \quad (3.3)$$

*isto é indica se os valores de  $X_i$  estão próximos ou distantes da média  $\mu_i$ , logo valores grandes  $\mu_i$  indicam uma maior dispersão de valores em relação à media de distribuição.*

**Definição 5** (Covariância). *A covariância entre os valores de  $i$ -ésima e  $j$ -ésima variáveis de vetor  $X$  é definida por:*

$$Cov(X_i, X_j) = \mu_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)] \quad (3.4)$$

*A partir do cálculo da covariância podemos medir o grau de relacionamento linear entre duas variáveis entre duas variáveis aleatória, embora a covariância tenha informação sobre o relacionamento linear entre duas variáveis, é difícil julgar se a relação é forte ou não, já que não tem valores de referência, máximo e mínimos, entretanto quando se tem muitos valores de covariância é aconselhavel guardar em uma matriz.*

**Definição 6** (Matriz de covariância). *A matriz de variâncias e covariâncias do*

vetor aleatório  $X$  é denotado por:

$$\text{Cov}(X) = V(X) = \text{Var}(X) = \Sigma_{p \times p} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

A matriz de covariância é simétrica, ou seja,  $\sigma_{12} = \sigma_{21}$

**Definição 7** (Correlação). O coeficiente de correlação entre as  $i$ -ésima e  $j$ -ésima variáveis de vetor  $X$  é definido por

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}} = \frac{\sigma_{ij}}{\sigma_i\sigma_j} \quad (3.5)$$

$\rho_{ij}$  varia no intervalo  $-1$  a  $1$ ,  $i, j = 1, 2, \dots, p$ . Quando  $i = j$ , a expressão anterior torna-se igual a  $1$ . A correlação é uma medida mais adequada para avaliar o grau de relacionamento linear entre as duas variáveis quantitativas, quando o  $\rho_{ij}$  está próximo de  $1$  significa que o relacionamento linear crescente positivo (crescimento), se estiver próximo de  $-1$  tem um relacionamento linear negativo (decréscimo), se estiver próxima de zero quer dizer que não há relacionamento linear. Porém se há muitos coeficientes de correlação fica mais viável guarda em uma matriz denominada de matriz de correlação

**Definição 8** (Matriz de Correlação). A matriz de correlação do vetor aleatório  $X$  é denotada por:

$$P_{p \times p} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \rho_{23} & \cdots & \rho_{2p} \\ \rho_{31} & \rho_{32} & 1 & \cdots & \rho_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \rho_{p3} & \cdots & 1 \end{bmatrix}$$

**Definição 9** (Matrizes de covariâncias e correlação de dois vetores aleatórios). Quando há presença de dois vetores aleatórios  $X$  e  $Y$ , tanto é possível calcular a matriz de covariância e de correlação de cada vetor, como calcular a matriz de covariância e correlação entre os dois vetores  $X$  e  $Y$ , onde essas matrizes serão denominadas por:  $\Sigma_x, \Sigma_y$  e  $\Sigma_{xy}$  e  $P_x, P_y$  e  $P_{xy}$ , respectivamente.

$$\Sigma_{xy} = \begin{bmatrix} \text{cov}(X_1, Y_1) & \text{cov}(X_1, Y_2) & \cdots & \text{cov}(X_1, Y_p) \\ \text{cov}(X_2, Y_1) & \text{cov}(X_2, Y_2) & \cdots & \text{cov}(X_2, Y_p) \\ \text{cov}(X_3, Y_1) & \text{cov}(X_3, Y_2) & \cdots & \text{cov}(X_3, Y_p) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_p, Y_1) & \text{cov}(X_p, Y_2) & \cdots & \text{cov}(X_p, Y_p) \end{bmatrix}$$

A matriz de covariância entre  $X$  e  $Y$  é a matriz  $\Sigma_{xy}$

$$P_{pxp} = \begin{bmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \rho_{23} & \cdots & \rho_{2p} \\ \rho_{31} & \rho_{32} & \rho_{33} & \cdots & \rho_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \rho_{p3} & \cdots & \rho_{pp} \end{bmatrix}$$

sendo que

$$\rho_{ij} = \frac{\text{cov}(X_{ij})}{\sqrt{\text{cov}(X_{ii})\text{cov}(X_{jj})}} \quad (3.6)$$

**Definição 10** (Variância Total). A variância total do vetor aleatório  $X$  é definido como traço ( $\Sigma_{pxp}$ ) =  $\text{tr}(\Sigma_{pxp}) = \rho_{11} + \rho_{22} + \dots + \rho_{pp}$

O traço da matriz sintetiza a variância global da distribuição multivariada, e com isso pode-se analisar o quanto os dados estão dispersos no sentido global.

**Definição 11** (Variância Generalizada). A variância generalizada de um vetor  $X$  é obtida através do cálculo do determinante da matriz  $|\Sigma_{xy}|$ , logo o desvio-padrão é  $|\Sigma_{xy}|^{\frac{1}{2}}$ . A Variância Generalizada informa também a dispersão dos dados, entretanto ela é influenciada pelas covariâncias (ou correlações) entre as variáveis



Álgebra linear é de fundamental importância tendo em vista que o objetivo é reduzir o número de variáveis sem perda informações, ou seja resumir a informação e para que seja tal procedimento logre êxito faz-se necessário o conhecimento em álgebra, principalmente em Combinações Lineares e Transformação Linear que serão vistos também a seguir de forma superficial

## Álgebra Linear

**Definição 12** (Combinação Linear). *Segundo (STRANG, 2012) Seja um espaço vetorial real  $V$  (ou complexo),  $v_1, v_2, \dots, v_n$  e  $a_1, \dots, a_n$  números reais. Logo o vetor:*

$$v = a_1v_1 + a_2v_2 + \dots + a_nv_n$$

*é um elemento de  $V$  é um elemento ao que chamamos combinação linear de  $v_1, v_2, \dots, v_n$*

*Uma vez fixado vetores  $v_1, \dots, v_n$  em  $V$ , o conjunto  $W$  de todos os vetores de  $V$  que são combinação linear destes, é um subespaço vetorial.  $W$  é chamado subespaço gerado por  $v_1, \dots, v_n$  tal processo é descrito, (BOLDRIN, 1986):*

$$W = [v_1, \dots, v_n]$$

ou

$$W = [v_1, \dots, v_n] = \{v \in V; v = a_1v_1 + \dots + a_nv_n, a_i \in \mathbb{R}, 1 \leq i \leq n\}$$

Uma outra caracterização de subespaço gerado é a seguinte:  $W = [v_1, \dots, v_n]$  é o menor subespaço de  $V$  que conjunto de vetores  $\{v_1, \dots, v_n\}$ , no sentido de que qualquer outro subespaço  $W'$  de  $V$  que contenha  $\{v_1, \dots, v_n\}$  satisfará  $W' \supset W$ .

Tanto álgebra linear como na estatística é importante saber se um vetor é uma combinação linear de outro vetores

**Definição 13** (Dependência e Independência Linear). *Seja  $V$  um espaço vetorial e  $v_1, \dots, v_n \in V$ . Dizemos que o conjunto  $\{v_1, \dots, v_n\}$  é linearmente independente (LI), ou que os vetores  $v_1, \dots, v_n$  são **LI**, se a equação*

$$a_1v_1 + \dots + a_nv_n = 0$$

,

implica que  $a_1 = a_2 = \dots = a_n$ . No caso em que exista algum  $a_i \neq 0$  dizemos que  $\{v_1, \dots, v_n\}$  é *linearmente dependente* (LD), ou que os vetores. Vetores linearmente dependentes podem ser caracterizados de uma outra maneira.

**Teorema 3.1.1.**  $\{v_1, \dots, v_n\}$  é LD se, e somente se um destes vetores for uma combinação linear de outros

**Definição 14** (Transformação Linear). *Sejam  $V$  e  $W$  dois espaço vetoriais. Uma transformação linear é uma função de  $V$  em  $W$ ,  $F : V \rightarrow W$ , que satisfaz as seguintes condições:*

*i. Quaisquer que sejam  $\mathbf{u}$  e  $\mathbf{v}$*

$$F(\mathbf{u} + \mathbf{v}) = F(\mathbf{u}) + F(\mathbf{v})$$

*ii. Quaisquer que sejam  $k \in \mathbb{R}$  e  $\mathbf{v} \in V$*

$$F(k\mathbf{v}) = kF(\mathbf{v})$$

**Definição 15** (Autovalor e Autovetor). *De acordo com (STRANG, 2012) **autovetor** de uma matriz  $\mathbf{A}_{n \times n}$ , é um vetor não-nulo  $\mathbf{x}$  tal que  $\mathbf{Ax} = \lambda\mathbf{x}$  para algum escalar  $\lambda$ . Um escalar  $\lambda$  é chamado de **autovalor** para  $\mathbf{A}$  se existe solução não-trivial  $\mathbf{x}$  para  $\mathbf{Ax} = \lambda\mathbf{x}$ , este  $\mathbf{x}$  é chamado de autovetor associado a  $\lambda$*

## Decomposição Espectral

De acordo com (MINGOTI, 2005), a decomposição de matrizes de covariância e de correlação em seus autovalores e autovetores normalizados, um teorema de suma importância em estatística multivariada e a da decomposição espectral.

**Teorema 3.1.2.** *Seja  $\Sigma_{p \times p}$  uma matriz de covariância. Então existe uma matriz*

ortogonal  $O_{p \times p}$ , isto é,  $O'O = OO' = I_{p \times p}$

$$O'\Sigma O = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_p \end{bmatrix} = \Lambda$$

onde  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$  são autovalores ordenados de forma decrescente na matriz  $\Sigma_{p \times p}$ , no caso da decomposição espectral considera a matriz  $\Sigma_{p \times p}$  é similar à matriz  $\Lambda_{p \times p}$ , portanto diz-se que:

i.

$$\det(\Sigma_{p \times p}) = |\Sigma_{p \times p}| = |\Lambda_{p \times p}| = \prod_{i=1}^p \lambda_i$$

ii.

$$\text{traco}(\Sigma_{p \times p}) = \text{traco}(\Lambda_{p \times p}) = \lambda_1 + \lambda_2 + \dots + \lambda_p$$

A  $i$ -ésima coluna da matriz  $O_{p \times p}$  é o autovetor normalizado  $e_i$  correspondete ao autovalor  $\lambda_i$ ,  $i = 1, 2, 3, \dots, p$ , que é agrupado por:

$$e_i = \begin{bmatrix} e_{i1} \\ e_{i2} \\ \vdots \\ e_{ip} \end{bmatrix}$$

A matriz  $O$  é dada por  $O = [e_1 e_2 \dots e_p]$  e pelo teorema da decomposição espectral tem-se que a seguinte igualdade é valida:

$$\Sigma_{p \times p} = O \Lambda O' = \sum_{i=1}^p \lambda_i e_i e_i' \quad (3.7)$$

onde  $e_i$  um vetor de comprimento igual a 1, isto é

$$\| e_i \| = (e_{i1}^2 + e_{i2}^2 + \dots + e_{ip}^2)^{1/2} = 1$$

## 3.2 Análise das Componentes Principais

Segundo (MINGOTI, 2005) análise das componentes principais é uma abordagem estatística dos fatores comuns, é uma abordagem estatística que pode ser usada para analisar interrelações entre um grande número de variáveis e explicar essas variáveis em termos de fatores. O objetivo é encontrar um meio de condensar a informação contida em um número de variáveis originais em um conjunto menor de fatores com a perda mínima de informação. Pelo fato de fornecer uma estimativa empírica da "estrutura" das variáveis consideradas.

### 3.2.1 Componentes Principais Extraída da Matriz de Covariância

Seja  $X = (X_1, X_2 \dots X_p)'$  um vetor aleatório com um vetor de médias  $\mu = (\mu_1, \mu_2 \dots \mu_p)'$  e matriz de covariância  $\Sigma_{p \times p}$ . Sejam  $\lambda_1 \geq \lambda_2 \dots \geq \lambda_p$  os autovalores da matriz  $\Sigma_{p \times p}$ , com os respectivos autovetores normalizados  $e_1, e_2, \dots, e_p$ , isto é, os autovetores  $e_i$  satisfazem as seguintes condições:

- i.  $e_i e_j = 0$  para todo  $i \neq j$ ;
- ii.  $e_i e_i = 1$  para todo  $i = 1, 2, \dots, p$ ;
- iii.  $\Sigma_{p \times p} e_i = \lambda_i e_i$  para todo  $i = 1, 2, \dots, p$

Considere o vetor  $Y$  que surge através da combinação linear entre a matriz ortogonal de autovetores  $O_{p \times p}$ , ou seja

$$\begin{bmatrix} e_{11} & e_{12} & \cdot & \cdot & \cdot & e_{p1} \\ e_{12} & e_{22} & \cdot & \cdot & \cdot & e_{p2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ e_{1p} & e_{2p} & \cdot & \cdot & \cdot & e_{pp} \end{bmatrix}$$

,

o vetor média também pode ser representado como a multiplicação do vetor  $\mu$  por  $O'$ , a matriz de covariância passa a ser uma matriz diagonal  $\Lambda_{p \times p}$ , e os autovalores

estão localizados na diagonal principal, de ordem decrescente, com isso as novas variáveis não estão mais correlacionadas, facilitando a interpretação conjunta, essa é a idéia do PCA usar combinação lineares em  $Y$ , para diminuir as quantidades de variáveis que antes era  $p$  e passa ser  $k$ , sendo  $k < P$ , contudo a variância total e a generalizada são iguais dos vetores  $X$  e  $Y$ , seguem abaixo algumas definições importantes:

**Definição 16** (Componentes Principais). *A componente principal da matriz de covariância  $\Sigma_{p \times p}$  na posição  $j$ -ésima posição é definida como:*

$$Y_j = e_j'X = e_{j1}X_1 + e_{j2}X_2 + \dots + e_{jp}X_p \quad (3.8)$$

A média e a variância de  $Y_j$  podem seguir a mesma ideia, respectivamente

$$E[Y_j] = e_j'\mu = e_{j1}\mu_1 + e_{j2}\mu_2 + \dots + e_{jp}\mu_p$$

$$Var[Y_j] = e_j'\Sigma_{p \times p}e_j = \lambda_j$$

Sendo que  $Cov[Y_J, Y_K] = 0, j \neq k$ . A variância de cada componente principal é representado pelo autovalor de mesma posição, ou seja,  $\lambda_j$  corresponde a variância de  $Y_j$ . Entretanto os autovalores estão ordem decrescente.

**Definição 17** (Variância Total). *O quanto é explicado da variância total de  $X$  pela  $j$ -ésima componente principal é calculado através da fórmula:*

$$\frac{Var[Y_j]}{VariânciaTotaldeX} = \frac{\lambda_j}{Traco(\Sigma_{p \times p})} = \frac{\lambda_j}{\sum_{i=1}^p \lambda_i} \quad (3.9)$$

**Definição 18** (Variância Explicada). *As  $k$  primeiras componentes principais explicam a variância total, em proporções, e são definidas como:*

$$\frac{\sum_{j=1}^k Var[Y_j]}{VariânciaTotaldeX} = \frac{\sum_{j=1}^k \lambda_j}{Traco(\Sigma_{p \times p})} = \frac{\sum_{j=1}^k \lambda_j}{\sum_{i=1}^p \lambda_i} \quad (3.10)$$

Se as  $k$  primeira componentes explicam grande parte da variância total do vetor  $X$ , pode se restringir apenas as  $k$  componentes, facilitando o trabalho já que houve uma diminuição e sem que haja muita perda de informações.

### Estimação das Componentes Principais: Matriz de Covariância

Anteriormente partimos do pressuposto de que a matriz de covariância não precisa ser estimada, porém na maioria dos casos acontece de modo contrário, ou seja, temos que estimá-la através dos dados amostrais  $S_{p \times p}$ .

Sejam  $\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_p$  os autovalores da matriz  $S_{p \times p}$  e sejam  $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_p$  os correspondentes autovetores normalizados. Logo a  $j$ -ésima componente principal amostral estimada é definida por:

$$\hat{Y}_j = \hat{e}'_j X = \hat{e}'_{j1} X_1 + \hat{e}'_{j2} X_2 + \dots + \hat{e}'_{jp} X_p, j = 1, 2, \dots, p$$

**Propriedade 1.** A variância estimada de  $\hat{Y}_j$  é igual a  $\hat{\lambda}_j$ ,  $j=1,2,\dots,p$ .

**Propriedade 2.** Não existe covariância entre as componentes  $\hat{Y}_1$  e  $\hat{Y}_k$ , portanto elas não estão mais correlacionadas.

**Propriedade 3.** Para calcular o quanto é explicado pela nova componente principal

$$\frac{Var[\hat{Y}_j]}{VariânciaTotalEstimada de X} = \frac{\hat{\lambda}_j}{traco(S_{p \times p})} = \frac{\hat{\lambda}_j}{\sum_{i=1}^p \hat{\lambda}_i}$$

**Propriedade 4.** A correlação entre a componente principal, feita a partir da amostra  $S_{p \times p}$  e a variável aleatória  $X_i$ ,  $i=1,2,\dots,p$  é dada por:

$$r_{\hat{Y}_j, X_i} = \frac{\hat{e}_{ji} \sqrt{\hat{\lambda}_j}}{\sqrt{s_{ii}}}$$

**Propriedade 5.** Segundo o teorema da decomposição espectral, a matriz de covariância,  $S_{p \times p}$ , pode ser escrita da seguinte forma:

$$S_{p \times p} = \sum_{j=1}^p \hat{\lambda}_j \hat{e}_j \hat{e}'_j$$

fazendo analogia para as  $k$  componentes estimadas temos:

$$S_{p \times p} \approx \sum_{j=1}^k \hat{\lambda}_j \hat{e}_j \hat{e}_j'$$

Ao usarmos métodos de obtenção das componentes principais através da matriz de covariância corremos o risco de ter uma discrepância entre os valores da variância, muita das vezes por questões de medidas, então costuma-se a normalizar os dados, uma das formas de normalizar é pela sua média e desvio-padrão, logo as componentes serão retirados da matriz de covariância, cujo dados originais já foram normalizados, esse procedimento é similar a obter as componentes pela matriz de correlação  $P_{p \times p}$  das variáveis originais  $X_i$

### Estimação das Componentes Principais: Matriz de Correlação

Suponhamos que  $Z_i = (X_i - \mu_i)/\sigma_i$  e  $Var(X_i) = \sigma_i^2$ ,  $i=1,2,\dots,p$ . A matriz de covariância das variáveis normalizadas,  $Z_i$ , será  $P_{p \times p}$ . E será feito o mesmo procedimento da estimação de componentes principais pela matriz de covariância. Logo a componente principal da matrix  $P_{p \times p}$  será definida por:

$$Y_j = e_j' Z = e_{j1} Z_1 + e_{j2} Z_2 + \dots + e_{jp} Z_p. \quad (3.11)$$

sendo que:

- (a) A variância de  $Y_j$  é igual a  $\lambda_j$  e a covariância entre  $Y_j$  e  $Y_k$  é igual a zero, para qualquer  $j \neq k$ ;
- (b) A correlação entre a componente  $Y_j$  e a variável padronizada  $Z_i$  é igual a:

$$\mathbf{r}_{Y_j, Z_i} = \mathbf{r}_{Y_j, X_i} = e_{ji} \sqrt{\lambda_j}$$

- (c) A variância total do vetor aleatório  $Z = (Z_1, Z_2, \dots, Z_p)'$  é igual ao traço da matriz  $P_{p \times p}$ , que é igual ao valor  $p$ , ou seja, o número de variáveis medidas em cada elemento amostral, logo a proporção da variância total explicada pela  $j$ -ésima componente é igual a  $\frac{\lambda_j}{p}$ ,  $j = 1, 2, \dots, p$ .

Vale ressaltar que o número obtido na matriz de covariância usando estimação de componentes principais pela matriz de covariância é diferente da estimação de

componentes principais pela matriz de correlação. A principal diferença entre os dois métodos, é que no caso da análise dessa matriz, a explicação da matriz de covariância está concentrada no primeiro componente, entretanto na correlação há uma melhor distribuição da explicação.

### 3.2.2 Critérios para Determinação do Número $k$ de Componentes Principais

Como o objetivo desta análise é reduzir a dimensionalidade do espaço amostral, ou seja, resumo de informações onde se precisa de  $p$ -variáveis e passará a  $k$  variáveis,  $k < p$ , portanto faz-se necessário ter alguns critérios de escolha na quantidade de componentes, há três procedimentos muito indicadas, sendo que duas matemáticas e a outra seria o critério do pesquisador.

#### Procedimento: Análise de representatividade em relação à variância total

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{j=1}^p \hat{\lambda}_j} = \gamma. \quad (3.12)$$

Não existe um valor definido para  $\gamma$  e sua escolha deverá ser feita de acordo com a natureza dos dados. Ao utilizar a técnica de componentes principais, há uma redução de  $p$ -variáveis para  $k$ -variáveis, de um forma bem simples, logo a utilidade prática das componentes decresce com o crescimento do número de componentes tornando difícil a interpretação dos dados, há situações principalmente em dados geofísico que o máximo que pode o valor  $\gamma$  é 90%. Porém existem divergência na análise de  $\gamma$ , pois como nós vimos há dois métodos da estimação de componentes pela matriz de covariância e pela matriz de correlação, quando as componentes principais são extraídas da matriz de correlação, necessita-se de um número maior de componentes para se alcançar o  $\gamma$ , se comparado ao  $\gamma$  extraído através da matriz de covariância. O uso da matriz de correlação para estimação da componentes principais faz uso de um critério que ajuda na escolha de  $k$  chamado de critério de Kaiser e diz que: *apenas manter no sistema apenas as componentes relacionadas àqueles autovalores  $\hat{\lambda}_i \geq 1$ , ou seja, mantém as combinações lineares que conseguem explicar pelo menos a quantidade de variância de uma variável original padronizada. Esse critério também pode ser aplicado quando a análise é feita com a matriz de covariâncias, pode-se manter no sistema as componentes relacionadas aos autovalores que são maiores ou*



*iguais a  $\hat{\lambda}_m$ .*

$$\hat{\lambda}_m = \frac{\sum_{j=1}^p \hat{\lambda}_j}{p} \quad (3.13)$$

# Decomposições Tensoriais

## 4.1 Introdução

---

Existem várias abordagens na literatura para a definição do termo tensor. Elas geralmente dependem de um domínio científico no qual eles são usados. No caso em geral, o tensor é definido em um sistema de coordenadas generalizado onde os eixos de coordenadas são curvas e não necessariamente ortogonais. Neste contexto, o tensor é tratado como uma entidade matemática que desfruta a propriedade multilinear depois uma mudança de coordenadas. Um tensor de  $N$ -ésima ordem é interpretado aqui como uma matriz que possui uma dependência linear com respeito a  $N$ - espaço vetoriais. Tensores são também são usados como um sinônimo de *matrizes multidimensionais*, também conhecidos como uma *multi-way arrays*. Em casos especiais como um tensor de ordem 2 é uma matriz, um tensor de ordem 1 é um vetor e um tensor de ordem 0 é um escalar. Provado que um tensor é uma forma multilinear e tem o seu próprio espaço linear associado, operações lineares comuns que são válidas para matrizes que podem ser estendidas para altas ordens.

## 4.2 Por Que Utilizar Decomposições Tensoriais?

---

De acordo com (CICHOCKI et al., 2009) as fatorações matriciais como *Principal Component Analysis* (PCA) e *Singular Value Decomposition* (SVD) são ferramentas importantes para a redução de dimensionalidade, redução de ruído e mineração de dados. No entanto essas fatorações possuem apenas representação em duas dimensões, como por exemplo, espaço e tempo tornando sua utilização limitada. Em vários estudos as estruturas de dados analisadas possuem na maioria das vezes

ordem superior a dois, por exemplo, ensaios, condições de trabalho, indivíduos e grupos, dimensões intrínsecas de espaço, tempo e frequência.

Dessa forma se observarmos o mundo em apenas duas dimensões utilizando fatorações matriciais podemos obter resultados inadequados para um determinado conjunto de dados analisado, para evitar essa incoerência dos resultados obtidos devemos utilizar técnicas adequadas de decomposição tensorial. As decomposições tensoriais mais utilizadas são os modelos de Tucker e o modelo PARAFAC, esse é mais simples que os modelos Tucker, pois apresenta uma fácil interpretação dos resultados obtidos.

### 4.3 Álgebra tensorial Básica

**Definição 19** (Notação escalar). *De acordo com (KOLDA; BADER, 2009) Seja  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$  é um tensor de  $N$ -ésima ordem. Um componente de  $X$  é especificado como*

$$x_{i_1, i_2, \dots, i_N} = [\mathcal{X}]_{i_1, i_2, \dots, i_N} \quad (4.1)$$

onde  $i_n$  é a  $n$ -ésima dimensão de  $\mathcal{X}$ , também chamada de modo- $n$  de  $\mathcal{X}$

**Definição 20** (Produto Interno). *(STRANG, 2012) afirma que o produto interno de dois tensores  $\mathcal{X}$  e  $\mathcal{Y}$  de mesma ordem  $N$  é dado por:*

$$\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \dots \sum_{i_N=1}^{I_N} x_{i_1, i_2, \dots, i_N} y_{i_1, i_2, \dots, i_N} \quad (4.2)$$

*Caso similar ao das matrizes, onde a noção de ortogonalidade entre dois tensores é análogo no produto interno,  $\mathcal{X}$  e  $\mathcal{Y}$  são ditos mutuamente tensores ortogonais se (2) for igual a 0.*

**Definição 21** (Produto Externo). *Seja  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$  e  $\mathcal{Y} \in \mathbb{C}^{J_1 \times J_2 \times \dots \times J_N}$  seja dois tensores de ordem  $N$  e  $M$ , respectivamente. O produto externo entre  $\mathcal{X}$  e  $\mathcal{Y}$  é dado por:*

$$[\mathcal{X} \circ \mathcal{Y}]_{i_1, i_2, \dots, i_N, j_1, j_2, \dots, j_N} = x_{i_1, i_2, \dots, i_N} y_{j_1, j_2, \dots, j_N} \quad (4.3)$$

O produto externo de dois tensores forma outro tensor cuja ordem é dada pela soma das ordens dos outros dois tensores. A equação (3) é uma generalização do conceito do produto externo de dois vetores, que gera uma matriz (tensor de segunda ordem). Também é dado uma noção de rank-1, como um caso especial.

**Definição 22** (Tensor de rank-1). Segundo (KOLDA; BADER, 2009), um tensor de rank-1  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$  é o tensor que pode ser escrito como o produto externo de vetores  $\mathbf{u}^{(1)} \in \mathbb{C}^{I_1}$ ,  $\mathbf{u}^{(2)} \in \mathbb{C}^{I_2}$ ,  $\dots$ ,  $\mathbf{u}^{(N)} \in \mathbb{C}^{I_N}$

$$x_{i_1, i_2, \dots, i_N} = u_{i_1}^{(1)} u_{i_2}^{(2)} \dots u_{i_N}^{(N)} \quad (4.4)$$

Os vetores  $u_n$  são chamadas de componentes de  $\mathcal{X}$ . Como um caso especial, a matriz de rank-1 é dada pelo produto externo de dois vetores. Como será mostrado a posteriormente, decomposições tensoriais são em geral combinações lineares de tensores de rank-1.

**Definição 23** (Rank). (SMILDE; BRO; GELADI, 2004) O rank de um tensor arbitrário  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$  é representado por  $R = r(\mathcal{X})$ , é o número mínimo de tensores de rank-1 que produz o tensor  $\mathcal{X}$  com combinações lineares.

**Definição 24** (Norma de Frobenius). De acordo (ALMEIDA, 2007) a norma de Frobenius de um tensor  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$  é definido como:

$$\|\mathcal{X}\|_F = \sqrt{\langle \mathcal{X}, \mathcal{X} \rangle} = \left( \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \dots \sum_{i_N=1}^{I_N} |x_{i_1, i_2, \dots, i_N}|^2 \right)^{1/2} \quad (4.5)$$

A norma de Frobenius pode ser interpretada como a medida da "energia" no tensor.

**Definição 25** (Desdobramento matricial). (SMILDE; BRO; GELADI, 2004) a matriz desdobrada  $\mathbf{X}_n$  de um tensor de ordem- $n$   $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$  é definida como uma matriz  $I_n \times I_1 I_2 \dots I_{n-1} I_{n+1} I_N$ , cujas colunas são vetores de dimensão- $I_n$  obtido pela variação do índice  $i_n$  e mantendo os outros índices fixos. A ordem de apresentação dos vetores de modo- $n$  pode variar de uma definição para outra. Uma visualização de três representações de desdobramento  $\mathbf{X}_1$ ,  $\mathbf{X}_2$  e  $\mathbf{X}_3$  para um tensor  $\mathcal{X}$  de terceira ordem  $\in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$  é mostrado na figura 1

**Definição 26** (Rank modo- $n$ ). Seja  $\mathbf{X}_n$  um desdobramento matricial de modo- $n$  do tensor  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ . O rank do tensor  $\mathcal{X}$  de modo- $n$  é a dimensão do espaço

vetorial gerado por vetores de modo- $n$  (ou seja, as colunas de  $X_n$ ). A definição de rank de modo- $n$  é uma generalização do conceito clássico de rank para matrizes. Contrária ao casos da matriz, onde  $R_1 = R_2 = R$  (ou seja rank linha é rank coluna), ranks de modo- $n$  de um tensor de alta ordem não são necessariamente os mesmos. Além disso quando os ranks de modo- $n$  são iguais, eles ainda podem ser diferentes da classificação do tensor. O rank de modo- $n$  é sempre inferior ou igual ao rank., exemplo  $R_n \leq R$ .

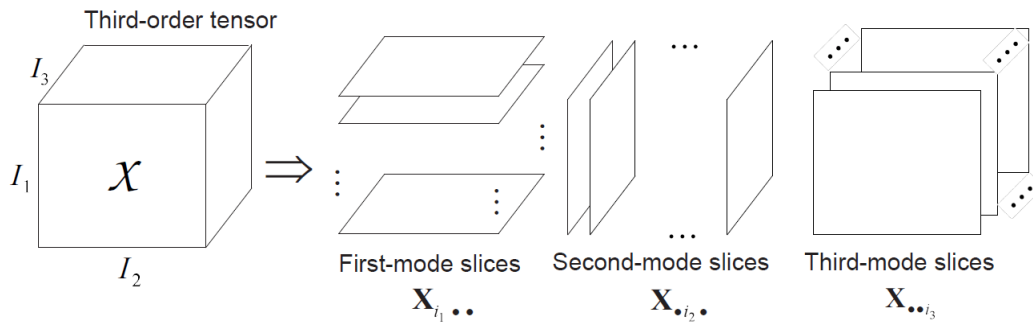


Figura 4.1: Desdobramento Matricial

**Definição 27** (Produto modo- $n$  ou Produto modal). O produto modo- $n$  de um tensor  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$  e uma matriz  $\mathbf{A} \in \mathbb{C}^{J_n \times I_n}$ , denotador por  $\mathcal{X} \bullet_n \mathbf{A}$  é especificado como:

$$[\mathcal{X}_n \bullet_n \mathbf{A}] =_{i_1, i_2, \dots, i_{n-1}, j_n, i_{n+1}, \dots, i_N} = \sum_{i_n=1}^{I_n} x_{i_1, i_2, \dots, i_{n-1}, i_n, i_{n+1}, \dots, i_N} a_{j_n, i_n} \quad (4.6)$$

O resultado de um produto modo- $n$  é um tensor de mesma ordem, mas com um novo  $n$ -ésima dimensão  $J_n$ .

O produto modo- $n$  é uma forma compacta de representação da transformação linear com tensores. É uma alternativa para a chamada convenção de somatório de Einstein. Esta notação faz clarear a analogia entre matriz e decomposição tensorial, bem como é dado uma intuitiva compreensão de decomposição tensorial.

## 4.4 A explicação em decomposição tensorial

Esta seção é focada na decomposição de tensores de alta ordem (multi-way arrays). A decomposição tensorial também é referida como *análise fatorial multi-way*

é uma área da álgebra multilinear que caracteriza um tensor como uma combinação linear do *produto externo de fatores*. Dependendo da abordagem considerada, a decomposição tensorial pode ser vista como uma generalização da Análise de Componentes Principais (PCA) ou da Decomposição de Valores Singulares (SVD) de ordens superiores a dois. A análise de um tensor em termos de seus *fatores decompostos* é útil no problema onde há uma *Mistura Multilinear* de diferentes fatores ou as contribuições devem ser identificadas para os dados medidos (ALMEIDA, 2007). É necessário salientar que as decomposições tensoriais podem ser generalizadas para N-ésima ordem, entretanto este trabalho aborda de maneira abrangente a ordem três, em alguns casos serão mostrados esse tipo de generalização.

### Decomposição Tucker-3

A decomposição Tucker-3 foi proposta por L. Tucker nos anos sessenta. Pode ser vista como uma extensão da análises de fatores bilinear para tensores de ordem três. A decomposição Tucker-3 é também um nome comum para denotar o Decomposição de Valores Singulares de Alta Ordem (HOSVD) de um tensor de terceira ordem. O modelo Tucker-3 é o que mais incorpora casos especiais de decomposições tensoriais.

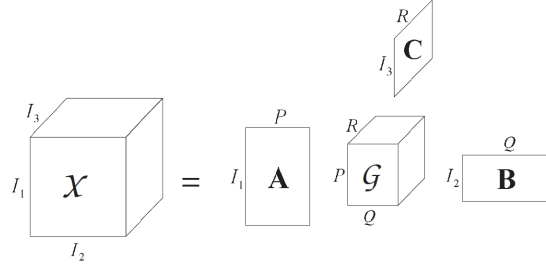
A decomposição tensorial de um tensor  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$  pode ser escrito na forma escalar como:

$$x_{i_1, i_2, i_3} = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R a_{i_1, p} b_{i_2, q} c_{i_3, r} g_{p, q, r} \quad (4.7)$$

onde  $a_{i_1, p} = [\mathbf{A}]_{i_1, p}$ ,  $b_{i_2, q} = [\mathbf{B}]$  e  $c_{i_3, r} = [\mathbf{C}]$ , são componentes escalar das

três matrizes decompostas  $[\mathbf{A}] \in \mathbb{C}^{I_1 \times P}$ ,  $[\mathbf{B}] \in \mathbb{C}^{I_2 \times Q}$  e  $[\mathbf{C}] \in \mathbb{C}^{I_3 \times R}$ , respectivamente, e  $g_{p, q, r} = [\mathcal{G}]_{p, q, r}$  é um componente escalar do *tensor núcleo*  $\mathcal{G} \in \mathbb{C}^{I \times P \times Q}$

A partir da equação (7), notamos que o tensor decomposto utilizando Tucker-3 é igual a uma combinação linear (ou soma ponderada) de Produtos Externos  $PQR$ , onde os coeficientes (ou ponderação) de cada termo do produto externo é correspondente a um componente escalar do tensor núcleo. Chamamos de  $P$  como o número de fatores no primeiro modo do tensor  $\mathcal{X}$ . Similarmente,  $Q$  e  $R$  denota o número de fatores no segundo e terceiro modo de  $\mathcal{X}$ . A decomposição Tucker-3 pode se referir como uma decomposição tensorial que permite *interações* entre fatores dos



**Figura 4.2:** Visualização da decomposição PARAFAC.

três modos do tensor. Uma ilustração da decomposição Tucker-3 é dada na figura 2.

A decomposição Tucker-3 pode também ser dita recorrendo à notação do produto modo- $n$  (modal) notação definida na equação (6).

$$\mathcal{X} = \mathcal{G} \bullet_1 \mathbf{A} \bullet_2 \mathbf{B} \bullet_3 \mathbf{C} \quad (4.8)$$

Alternativamente, podemos afirmar a decomposição Tucker-3 usando uma *notação de matriz-fatia*. Esta notação caracteriza um tensor por um conjunto de matriz-fatia que são obtida por "fatiamento" do tensor numa dada "direção". Cada matriz-fatia é obtida por fixação de um índice de um modo e a variação dos outros dois modos. Para um tensor de ordem três, há três possibilidade de direção de fatiamento. Nós chamamos  $\mathbf{X}_{i_1..} \in \mathbb{C}^{I_2 \times I_3}$  o  $i_1$ -ésimo *fatia do primeiro modo*,  $\mathbf{X}_{i_2..} \in \mathbb{C}^{I_3 \times I_1}$  o  $i_2$ -ésimo *fatia do segundo modo* e  $\mathbf{X}_{i_3..} \in \mathbb{C}^{I_1 \times I_2}$   $i_3$ -ésimo *fatia do terceiro modo*. Reescreveremos a decomposição Tucker-3 usando a notação de matriz-fatia, temos (7):

$$x_{i_1, i_2, i_3} = \sum_{q=1}^Q \sum_{r=1}^R b_{i_2, q} c_{i_3, r} \left( \sum_{p=1}^P a_{i_1, p} g_{p, q, r} \right) \quad (4.9)$$

e definir uma "equivalência" do núcleo (combinado ao primeiro modo) temos:

$$u_{i_1, q, r}^{(1)} = \sum_{p=1}^P a_{i_1, p} g_{p, q, r} = [\mathcal{G} \bullet_1 \mathbf{A}]_{i_1, q, r} \quad (4.10)$$

ou seja,  $u_{i_1, q, r}^{(1)} = [\mathcal{G} \bullet_1 \mathbf{A}]_{i_1, q, r}$ . A  $i$ -ésima matriz-fatia  $\mathbf{X}_{i_1..}$ ,  $i_1 = 1, \dots, I_1$ , é dado por:

$$\mathbf{X}_{i_1..} = \mathbf{B}\mathbf{U}_{i_1}^{(1)}\mathbf{C}^T, \quad i_1 = 1, \dots, I_1, \quad (4.11)$$

$\mathbf{U}_{i_1}^{(1)}$  é a  $i_1$ -ésima matriz-fatia do primeiro-modo da transformada do tensor núcleo  $\mathcal{U}^{(1)} \in \mathbb{C}^{I_1 \times Q \times R}$ . As notações das outras duas matriz-fatias são obtidas na forma similar, mudando a ordem do somatório na equação (7) e definição:

$$u_{p,i_2,r}^{(2)} = \sum_{q=1}^Q b_{i_2,q} g_{p,q,r}, \quad (4.12a)$$

$$u_{p,q,i_3}^{(3)} = \sum_{r=1}^R c_{i_3,r} g_{p,q,r} \quad (4.12b)$$

como componente escalar da transformada do tensor núcleo  $\mathcal{U}^{(2)} = [\mathcal{G} \bullet_2 \mathbf{B}] \in \mathbb{C}^{P \times I_2 \times R}$  e  $\mathcal{U}^{(3)} = [\mathcal{G} \bullet_3 \mathbf{C}] \in \mathbb{C}^{P \times Q \times I_3}$ . Isto nos leva a:

$$\mathbf{X}_{.i_2.} = \mathbf{C}\mathbf{U}_{i_2}^{(2)}\mathbf{A}^{(T)}, \quad i_2 = 1, \dots, I_2, \quad (4.13a)$$

e

$$\mathbf{X}_{..i_3} = \mathbf{A}\mathbf{U}_{i_3}^{(3)}\mathbf{A}^{(T)}, \quad i_3 = 1, \dots, I_3, \quad (4.13b)$$

Seja  $\mathbf{X}_1 \in \mathbb{C}^{I_3 I_1 \times I_2}$ ,  $\mathbf{X}_2 \in \mathbb{C}^{I_1 I_2 \times I_3}$  e  $\mathbf{X}_3 \in \mathbb{C}^{I_2 I_3 \times I_1}$  é o primeiro, segundo e terceiro modo das matrizes desdobradas de  $\mathbf{X}$ . Estas matrizes são definidas como:

$$\mathbf{X}_1 = \begin{bmatrix} \mathbf{X}_{..1} \\ \vdots \\ \mathbf{X}_{..I_3} \end{bmatrix}, \quad \mathbf{X}_2 = \begin{bmatrix} \mathbf{X}_{1..} \\ \vdots \\ \mathbf{X}_{I_1..} \end{bmatrix}, \quad \mathbf{X}_3 = \begin{bmatrix} \mathbf{X}_{.1.} \\ \vdots \\ \mathbf{X}_{.I_2.} \end{bmatrix}, \quad (4.14)$$

Elas podem ser mostradas a partir da equação (11), (13a) e (13b)  $\mathbf{X}_1$ ,  $\mathbf{X}_2$  e  $\mathbf{X}_3$  podem ser expressadas como:

$$\mathbf{X}_1 = (\mathbf{C} \otimes \mathbf{A})\mathbf{G}_1\mathbf{B}^T, \quad \mathbf{X}_2 = (\mathbf{A} \otimes \mathbf{B})\mathbf{G}_2\mathbf{C}^T, \quad \mathbf{X}_3 = (\mathbf{B} \otimes \mathbf{C})\mathbf{G}_3\mathbf{A}^T, \quad (4.15)$$

onde  $\mathbf{G}_1 \in \mathbb{C}^{RP \times Q}$ ,  $\mathbf{G}_2 \in \mathbb{C}^{PQ \times R}$  e  $\mathbf{G}_3 \in \mathbb{C}^{QR \times R}$  são matrizes desdobradas do tensor  $\mathcal{G}$ , no qual são construídas na mesma forma como (14)



$$\mathbf{G}_1 = \begin{bmatrix} \mathbf{G}_{..1} \\ \vdots \\ \mathbf{G}_{..R} \end{bmatrix}, \mathbf{G}_2 = \begin{bmatrix} \mathbf{G}_{1..} \\ \vdots \\ \mathbf{G}_{P..} \end{bmatrix}, \mathbf{G}_3 = \begin{bmatrix} \mathbf{G}_{.1.} \\ \vdots \\ \mathbf{G}_{.Q.} \end{bmatrix}, \quad (4.16)$$

a  $\otimes$  denota o produto de Kronecker. Cada um das três matrizes desdobradas (15) são diferentes rearranjos da mesma informação contida no tensor  $\mathcal{X}$ .

### Generalizando Tucker-3 para a $N$ -ésima ordem

A generalização da decomposição Tucker-3 de  $N$ -ésima ordem é simples. Vamos considerar um tensor  $\mathcal{X}$  de  $N$ -ésima ordem.  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ . Sua  $N$ -ésima ordem de decomposição Tucker pode ser expressada como:

$$x_{i_1, i_2, \dots, i_N} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \dots \sum_{r_N=1}^{R_N} a_{i_1, r_1}^{(1)} a_{i_2, r_2}^{(2)} \dots a_{i_N, r_N}^{(N)} g_{r_1, r_2, \dots, r_N}, \quad (4.17)$$

onde  $a_{i_n, r_n}^{(n)} = [\mathbf{A}^{(n)}]_{i_n, r_n}$  é um componente escalar do  $n$ -ésimo modo da matriz fator e  $g_{r_1, r_2, \dots, r_N} = [\mathcal{G}]_{r_1, r_2, \dots, r_N}$  é um componente escalar do  $N$ -ésimo ordem do tensor núcleo. A notação para o produto modo- $N$  ou produto modal pode ser escrito como:

$$\mathcal{X} = \mathcal{G} \bullet_1 \mathbf{A}^{(1)} \bullet_2 \mathbf{A}^{(2)} \dots \bullet_N \mathbf{A}^{(N)} \quad (4.18)$$

Voltemos ao caso de terceira ordem. A decomposição Tucker-3 não é única, uma vez que existem infinitas soluções para as matrizes decompostas e para o tensor núcleo levando ao mesmo tensor  $\mathcal{X}$ . Em outras palavras, a decomposição Tucker-3 permite transformações lineares arbitrárias sobre as três matrizes decompostas (provido que a inversa desta transformação é aplicada no tensor núcleo) sem afetar a reconstrução do tensor  $\mathcal{X}$ . Na ordem para ver este, nos definiremos matrizes não-singulares  $\mathbf{T}_a \in \mathbb{C}^{P \times P}$ ,  $\mathbf{T}_b \in \mathbb{C}^{Q \times Q}$  e  $\mathbf{T}_c \in \mathbb{C}^{R \times R}$ . Considerando a matriz decomposta, nós temos que:

$$\begin{aligned}\mathbf{X}_1 &= (\mathbf{C}\mathbf{T}_c\mathbf{T}_c^1 \otimes \mathbf{A}\mathbf{T}_a\mathbf{T}_a^{-1})\mathbf{G}_1(\mathbf{B}\mathbf{T}_b\mathbf{T}_b^{-1})^T \\ &= [(\mathbf{C}\mathbf{T}_c) \otimes (\mathbf{A}\mathbf{T}_a)][(\mathbf{T}_c^{-1} \otimes \mathbf{T}_a^{-1})\mathbf{G}_1\mathbf{T}_b^{-T}](\mathbf{B}\mathbf{T}_b)^T,\end{aligned}\quad (4.19)$$

$$\mathbf{X}_1 = (\tilde{\mathbf{C}} \otimes \tilde{\mathbf{A}})\tilde{\mathbf{G}}_1\tilde{\mathbf{B}}^T$$

onde  $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{T}_a$ ,  $\tilde{\mathbf{B}} = \mathbf{B}\mathbf{T}_b$  e  $\tilde{\mathbf{C}} = \mathbf{C}\mathbf{T}_c$  são matrizes decompostas transformadas e  $\tilde{\mathbf{G}}_1 = (\mathbf{T}_c^{-1} \otimes \mathbf{T}_a^{-1})\mathbf{G}_1\mathbf{T}_a^{-1}$  é tensor núcleo do Tucker-3 transformado. Na (19), aplicamos a seguinte propriedade do produto de Kronecker:

**Propriedade 1.1:** Dado  $\mathbf{A} \in \mathbb{C}^{I \times R}$ ,  $\mathbf{B} \in \mathbb{C}^{J \times S}$ ,  $\mathbf{C} \in \mathbb{C}^{R \times P}$  e  $\mathbf{D} \in \mathbb{C}^{S \times Q}$

$$\mathbf{AC} \otimes \mathbf{BD} = (\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) \quad (4.20)$$

Equação (19) diz que nós temos um número infinito de matrizes  $\tilde{\mathbf{A}}$ ,  $\tilde{\mathbf{B}}$ ,  $\tilde{\mathbf{C}}$  e  $\tilde{\mathbf{G}}_1$ , dando origem a mesma matriz  $\mathbf{X}_1$ . Esse fato indica claramente o geral da não-unicidade da decomposição Tucker-3. Unicidade completa das matrizes decompostas e o tensor núcleo da decomposição Tucker-3 é apenas possível em alguns casos especiais, onde pelo menos duas matrizes decompostas tenha alguma estrutura especial que permite uma determinação única das transformações das matrizes. Tem sido demonstrado que a unicidade parcial (ou seja, a unicidade de pelo menos alguns fatores) pode existir em caso onde o tensor núcleo Tucker-3 é restringido de ter vários elementos igual a zero.

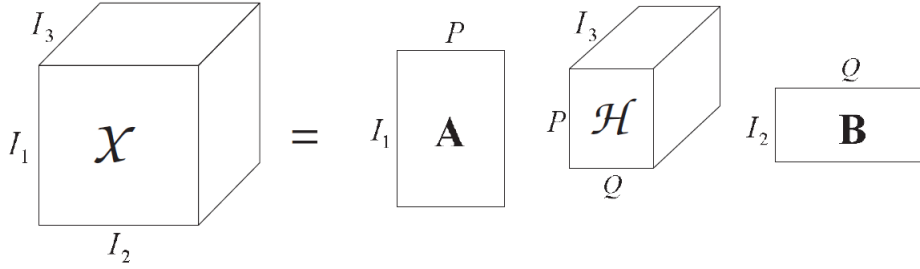
### Casos Especial: Tucker-2 e Tucker-1

Considere a decomposição Tucker-3 e (7) reescrito como:

$$x_{i_1, i_2, i_3} = \sum_{p=1}^P \sum_{q=1}^Q a_{i_1, p} b_{i_2, q} \left( \sum_{r=1}^R c_{i_3, r} g_{p, q, r} \right) = \sum_{p=1}^P \sum_{q=1}^Q a_{i_1, p} c_{i_2, q} h_{p, q, i_3}, \quad (4.21)$$

onde  $c_{i_3, r}$  foi absorvido pelo núcleo  $g_{p, q, r}$ , dando origem a um núcleo equivalente  $h_{p, q, i_3}$ , ou seja,  $h_{p, q, i_3} = [\mathcal{G} \bullet_3 \mathbf{C}]_{p, q, i_3}$ . Equação (21) é a notação escalar de uma equivalente decomposição Tucker-2. Note que a decomposição Tucker-2 é mais

simples do que o seu equivalente Tucker-3, pois o número de termos do produto externo foi reduzido para  $PQ$ . Uma decomposição Tucker-2 surge quando uma matriz, da decomposição Tucker-3, chamada de  $\mathbf{C}$ , é igual a uma matriz identidade. A notação para a fatia e o desdobramento para o caso do Tucker-2 poder ser facilmente obtida a partir da equação (11), (13), (14a e 14b) e (16) considerando  $\mathbf{C} = \mathbf{I}_{I_3}$  e  $\mathcal{G} = \mathcal{H} \in \mathbb{C}^{C \times Q \times I_3}$ . A decomposição Tucker-2 é ilustrada na Figura 3.



**Figura 4.3:** Visualização da decomposição Tucker-2.

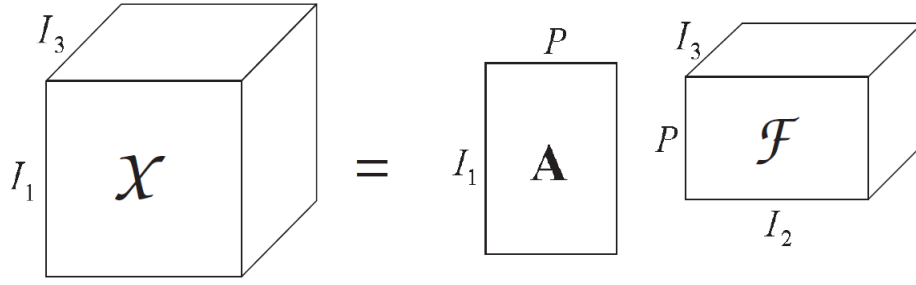
Agora, nos reescreveremos (7) como:

$$x_{i_1, i_2, i_3} = \sum_{p=1}^P a_{i_1, p} \left( \sum_{q=1}^Q \sum_{r=1}^R b_{i_2, q} c_{i_3, r} g_{p, q, r} \right) = \sum_{p=1}^P a_{i_1, p} f_{p, i_2, i_3}, \quad (4.22)$$

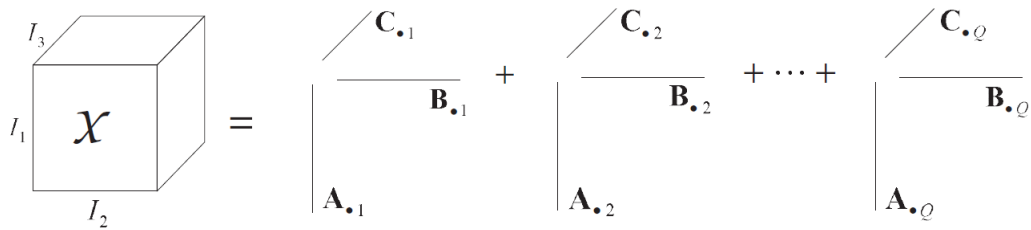
onde ambos  $b_{i_2, q}$  e  $c_{i_3, r}$  foi absorvido pelo núcleo  $g_{p, q, r}$ , resultando em outro núcleo  $f_{p, i_2, i_3}$ , ou seja,  $f_{p, i_2, i_3} = [\mathcal{G} \bullet_2 \mathbf{B} \bullet_3 \mathbf{C}]_{p, i_2, i_3}$ . Equação (22) é a notação escalar da decomposição Tucker-1. A decomposição Tucker-1 também surge a partir do Tucker-3 ocorre quando 2 matrizes, chamadas de  $\mathbf{B}$  e  $\mathbf{C}$  são iguais a matriz identidade. A notação para a fatia e o desdobramento para o modelo Tucker-1 é obtida a partir (11), (13), (14) e (16) fixando  $\mathbf{B} = \mathbf{I}_{I_2}$ ,  $\mathbf{C} = \mathbf{I}_{I_3}$  e  $\mathcal{G} = \mathcal{F} \in \mathbb{C}^{P \times I_2 \times I_3}$ . Figura 4 ilustra a decomposição Tucker-1.

### Decomposição Fatores Paralelos (PARAFAC)

Segundo (ALMEIDA, 2007) a decomposição FACTores PARAlelos (PARAFAC), também conhecida como DECOMPOSIÇÃO CANônica (CANDECOMP), foi independentemente desenvolvida por Hasham e Carol e Chang nos anos setentas. É também conhecida pela sigla CP (Candecomp-Parafac). Recentemente, esta decomposição foi usada em muitas aplicações para o processamento de sinais, dentre eles a aplicação e exploração na comunicação sem-fio



**Figura 4.4:** Visualização da decomposição Tucker-1.



**Figura 4.5:** Visualização da decomposição PARAFAC.

e entre outras aplicações. A decomposição PARAFAC também descreve a estrutura básica de cumulantes de ordem superior de dados multivariado em que todos métodos algébricos para Análise de Componente Independente (ICA).

Para um tensor terceira ordem, a decomposição de um tensor é um somatório do *produto Triplo* ou *Triad*. O PARAFAC pode ser afirmado como uma decomposição tensorial e a uma equivalência a uma soma de tensores de rank-1. A decomposição PARAFAC de um tensor  $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$  tem a seguinte notação escalar:

$$x_{i_1, i_2, i_3} = \sum_{q=1}^Q a_{i_1, q} b_{i_2, q} c_{i_3, q} \quad (4.23)$$

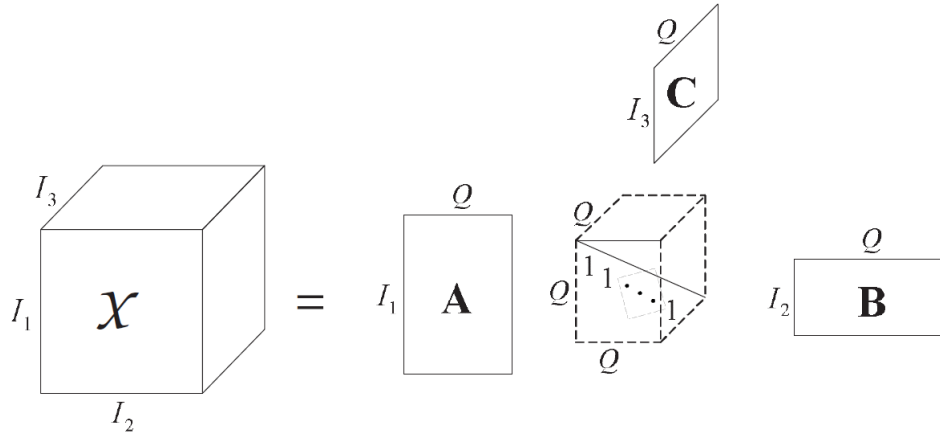
onde  $a_{i_1, q} = [\mathbf{A}]_{i_1, q}$ ,  $b_{i_2, q} = [\mathbf{B}]_{i_2, q}$  e  $c_{i_3, q} = [\mathbf{C}]_{i_3, q}$  são componentes escalar das matrizes decompostas  $\mathbf{A} \in \mathbb{C}^{I_1 \times Q}$ ,  $\mathbf{B} \in \mathbb{C}^{I_2 \times Q}$  e  $\mathbf{C} \in \mathbb{C}^{I_3 \times Q}$ , respectivamente.  $Q$  é o número de fatores, também conhecido como *rank* da decomposição. As colunas das primeira, segunda e terceira matriz decomposta  $\mathbf{A}$ ,  $\mathbf{B}$  e  $\mathbf{C}$  são chamadas, respectivamente, *Fator de Carregamento* de primeiro, segundo e terceiro modo. Outro sinônimo para as colunas de  $\mathbf{A}$ ,  $\mathbf{B}$  e  $\mathbf{C}$  são *padrões de carregamento* ou *vetores de carregamento*.

Usando notação do produto externo, a decomposição PARAFAC para um tensor

$\mathcal{X}$  de terceira ordem. Pode ser reescrito como:

$$\mathcal{X} = \sum_{q=1}^Q \mathbf{A}_{.q} \circ \mathbf{B}_{.q} \circ \mathbf{C}_{.q}.$$

Na figura 5, a decomposição PARAFAC de terceira ordem é visualizada como uma soma de  $Q$  tensores de  $rank$ -1. Uma visualização é retratada na Fig. 6, onde a decomposição PARAFAC é interpretada como um caso especial da decomposição Tucker-3 como um tensor núcleo identidade.



**Figura 4.6:** Visualização da decomposição PARAFAC é interpretada como um caso especial da decomposição Tucker-3

Das três possíveis direções de fatiamento de  $\mathcal{X}$ , obtemos a seguinte notação para a decomposição PARAFAC:

$$\begin{aligned} \mathbf{X}_{i_1..} &= \mathbf{B} D_{i_1}(\mathbf{A}) \mathbf{C}^T, \\ \mathbf{X}_{.i_2.} &= \mathbf{C} D_{i_2}(\mathbf{B}) \mathbf{A}^T, \\ \mathbf{X}_{..i_3} &= \mathbf{A} D_{i_3}(\mathbf{C}) \mathbf{B}^T, \end{aligned} \tag{4.24}$$

onde  $D_{i_1}(\mathbf{A})$  forma uma diagonal mantendo a  $i_1$ -ésima linha de  $\mathbf{A} \in \mathbb{C}^{I_1 \times Q}$  diagonal principal. Pelo empilhamento de linha-por-linha a primeira, segunda e a terceira da fatia do terceiro modo nós temos:

$$\begin{aligned}
\mathbf{X}_1 &= \begin{bmatrix} \mathbf{X}_{..1} \\ \vdots \\ \mathbf{X}_{..I_3} \end{bmatrix} = \begin{bmatrix} \mathbf{A}D_1(\mathbf{C}) \\ \vdots \\ \mathbf{A}D_{I_3}(\mathbf{C}) \end{bmatrix} & \mathbf{B}^T &= (\mathbf{C} \diamond \mathbf{A})\mathbf{B}^T \\
\mathbf{X}_2 &= \begin{bmatrix} \mathbf{X}_{1..} \\ \vdots \\ \mathbf{X}_{I_1..} \end{bmatrix} = \begin{bmatrix} \mathbf{B}D_1(\mathbf{A}) \\ \vdots \\ \mathbf{B}D_{I_1}(\mathbf{A}) \end{bmatrix} & \mathbf{C}^T &= (\mathbf{A} \diamond \mathbf{B})\mathbf{C}^T \\
\mathbf{X}_3 &= \begin{bmatrix} \mathbf{X}_{.1.} \\ \vdots \\ \mathbf{X}_{.I_2.} \end{bmatrix} = \begin{bmatrix} \mathbf{C}D_1(\mathbf{B}) \\ \vdots \\ \mathbf{C}D_{I_2}(\mathbf{B}) \end{bmatrix} & \mathbf{A}^T &= (\mathbf{B} \diamond \mathbf{C})\mathbf{A}^T \tag{4.25}
\end{aligned}$$

onde  $\diamond$  denota o Produto de Khatri-Rao (Matriz-Coluna Kronecker) Uma das propriedades mais interessantes do PARAFAC é a **unicidade**. Contrariamente nas decomposições bilinear (matriz), que em geral não são único para ranks superiores a um, (matrizes de rank-um até atingir um fator escalar), a decomposição PARAFAC de tensores de rank superiores a um, são únicos até a escala e permutação de fatores.

Os primeiros estudos de singularidade da decomposição PARAFAC, foram feitos nos anos setenta por Jennrich e Harshman. A prova mais profunda de singularidade foi fornecida por Kruskal. As condições *suficientes*, proveniente de Kruskal, para a unicidade da decomposição PARAFAC de tensores de terceira ordem de valores real. Por volta de duas décadas depois, Sidiropoulos *et al* (1984) estendeu as condições de Kruskal para a condição de tensores de valores complexo. Sidiropoulos & Bro generalizaram as condições de Kruskal, extendendo as condições de unicidade para tensores de ordem  $N$ . Ten Berge e Sidiropoulos (1983) mostraram que a condição de Kruskal não é apenas suficiente, mas também *necessária* para  $Q \in \{2, 3\}$ . Além disso, questões de unicidade do PARAFAC foram criadas por Jiang & Sidiropoulos (1985) das quais derivaram as condições suficientes e necessárias para unicidade do assim chamado modelo PARAFAC *restringido* (ou seja, quando pelo menos uma matriz, da decomposição, for rank-coluna completa). A mais acessível prova de unicidade foi provada usando a álgebra linear convencional. Um novo limite de singularidade que é mais abrangente do que o limite de Kruskal, foi derivado a partir de um link entre a decomposição PARAFAC e a decomposição simultaneamente de

matrizes.

O estudo das condições de unicidade do modelo PARAFAC é baseado em um conceito fundamental,  $k$ -rank (*Kruskal-rank*), muito mais restrito que o conceito usual de rank de matriz. O conceito de  $k$ -rank foi proposto por Kruskal em seu artigo semanal, foi usado primeiramente por Harshman e Lundy (1975). O conceito de  $k$ -rank foi usado como um conceito chave para declarar a unicidade no PARAFAC.

**Definição 28** ( $k$ -rank). *O rank de  $\mathbf{A} \in \mathbb{C}^{I_1 \times Q}$ , denotado por  $r_{\mathbf{A}}$ , é igual a  $r$  se  $\mathbf{A}$  contém pelo menos um conjunto de  $r$  de colunas linearmente independente, mas nenhum conjunto de  $r + 1$  colunas linearmente independente. O *Kruskal-rank* (ou  $k$ -rank) de  $\mathbf{A}$  é um número máximo  $k$  tal que todo conjunto de  $k$  colunas de  $\mathbf{A}$  é linearmente independente. Note que  $k$ -rank é sempre menor do que ou igual ao rank, sendo assim nós temos:*

$$k_{\mathbf{A}} \leq r_{\mathbf{A}} \leq \min(I_1, Q)$$

**Teorema 4.4.1.** *:Considere o conjunto de  $I_1$  matriz-fatias  $\mathbf{X}_{i_1 \dots} = \mathbf{B}D_{i_1}(\mathbf{A})\mathbf{C}^T$ ,  $i_1 = 1, \dots, I_1$ , definido em (1,25). Se*

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2Q + 2, \quad (4.26)$$

as matrizes  $\mathbf{A}$ ,  $\mathbf{B}$  e  $\mathbf{C}$  são únicas até a permutação e (complexo) dimensionamento de suas colunas [90,31]. Isto significa que qualquer matriz  $\tilde{\mathbf{A}}$ ,  $\tilde{\mathbf{B}}$  e  $\tilde{\mathbf{C}}$  satisfazendo a equação 1.25 são ligadas a  $\mathbf{A}$ ,  $\mathbf{B}$  e  $\mathbf{C}$  por:

$$\tilde{\mathbf{A}} = \mathbf{A}\mathbf{\Pi}\mathbf{\Delta}_1, \quad \tilde{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}\mathbf{\Delta}_2, \quad \tilde{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Delta}_3 \quad (4.27)$$

onde  $\mathbf{\Pi}$  é uma permutação de matriz e  $\mathbf{\Delta}_1$ ,  $\mathbf{\Delta}_2$  e  $\mathbf{\Delta}_3$  são matrizes diagonais satisfazendo a condição

$$\mathbf{\Delta}_1\mathbf{\Delta}_2\mathbf{\Delta}_3 = \mathbf{I}_Q \quad (4.28)$$

Condição (1.27) também é necessário se  $Q \leq 3$ . Contudo, ten Berge & Sidiropoulos (1983) provaram um simples contra-exemplo para o necessário de (1.27). Eles também afirmaram que a unicidade do PARAFAC depende de um caso particular e conjunta da presença de zeros nas matrizes obtidas na decomposição.

*Isto foi melhor explicado e esclarecido por Jiang & Sidiropoulos em (1985). Eles forneceram ao mesmo tempo condições necessárias e suficientes para a definição de unicidade para o PARAFAC, quando uma matriz obtida na decomposição possui rank-coluna completa.*

**Teorema 4.4.2.** *Para uso geral decomposição PARAFAC com  $Q > 1$ , duas condições são necessárias para a unicidade, são elas:*

$$\min(r_{\mathbf{A} \diamond \mathbf{B}}, r_{\mathbf{B} \diamond \mathbf{C}}, r_{\mathbf{C} \diamond \mathbf{A}}) = Q \quad e \quad \min(k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}}) \geq 2 \quad (4.29)$$

Condições (1.29) podem ser interpretadas na seguinte maneira. Note que  $(r_{\mathbf{A} \diamond \mathbf{B}}, \leq \min(I_1 I_2, Q), r_{\mathbf{B} \diamond \mathbf{C}} \leq \min(I_2 I_3, Q)$  e  $r_{\mathbf{C} \diamond \mathbf{A}} \leq \min(I_1 I_3, Q)$ . Note também que a maneira de  $k_{\mathbf{A}} \geq 2$ , é equivalente que a matriz  $\mathbf{A}$  não tem nenhuma coluna proporcional (Caso contrário  $k_{\mathbf{A}} = 1$ , de acordo com a definição de  $k$ -rank). É, portanto, equivalente afirma que, unicidade surge apenas se: i) *O produto de qualquer das duas dimensões do tensor é maior igual ao numeros de fatores e ii) nenhuma das três matrizes obtidas da decomposição pode ter um par colunas proporcionais.*

Ten Berge & Sidiropoulos mostraram que (1.29) é uma condição necessária e suficiente se uma matriz da decomposição é rank-coluna completo. Assumindo por exemplo, que  $\mathbf{C}$  é coluna-rank completo, ou seja,  $r_{\mathbf{C}} = Q$ , é facilmente checado que a condição (1.29), pode ser equivalentemente indicado como:

$$r_{\mathbf{A} \diamond \mathbf{B}} = Q \quad e \quad (\min k_{\mathbf{A}}, k_{\mathbf{B}}) \geq 2, \quad (4.30)$$

Significa dizer que o PARAFAC é único se e somente se o produto de Khatri-Rao  $\mathbf{A} \diamond \mathbf{B}$  é rank-coluna completa e ii) nem  $\mathbf{A}$  nem  $\mathbf{B}$  podem ter um par de colunas proporcionais. Na década de 80, foi também proposto uma condição necessária e equivalente, a qual é válida no caso geral da decomposição PARAFAC e é também mais fácil de verificar (1.29) e (1.30).

Como será mostrado mais adiante, a unicidade de pelo menos um subconjunto do fatores do PARAFAC é possível, mesmo em casos onde (um máximo de) duas matrizes fatores têm colunas proporcionais. Unicidade de pelo menos um subconjunto em um ou mais modos é chamado de (*Unicidade Parcial*). Foi primeiramente observado por Harshman, que salientou que a unicidade do PARAFAC pode ser quebrada em partes. Este conceito é útil para estudos de



propriedades de unicidades de modelos PARAFAC especiais, apresentando fatores proporcionais em um ou mais modos.

### PARAFAC de $N$ -ésima ordem

A decomposição PARAFAC de um tensor  $\mathcal{X} \in \mathbb{C}_{I_1 \times I_2 \times \dots \times I_N}$  pode ser declarados:

$$x_{i_1, i_2, \dots, i_N} = \sum_{q=1}^Q a_{i_1, q}^{(1)} a_{i_2, q}^{(2)} \dots a_{i_N, q}^{(N)} = \sum_{q=1}^Q \prod_{n=1}^N a_{i_n, q}^{(n)} \quad (4.31)$$

onde  $a_{i_n, q}^{(n)} = [\mathbf{A}^{(n)}]_{i_n, q}$ ,  $i_n = 1, \dots, I_n$ ,  $n = 1, \dots, N$ . Modelo (1.31) é uma soma de  $Q$ -fatores de  $N$ -produtos. Sidiropoulos & Bro proporcionaram uma condição de unicidade suficiente para o PARAFAC de ordem  $N$ .

$$\sum_{n=1}^N k_{\mathbf{A}^{(n)}} \geq 2Q + (N - 1) \quad (4.32)$$

Para o caso de ordem- $N$ , as condições necessárias seguem diretamente para os casos de terceira ordem, dado em (1.29). Contudo, a existência de uma condição necessária e suficiente para unicidade não é bem estabelecida na literatura.

### Algoritmo dos Mínimos Quadrados Alternados (ALS)

A estimação das três matrizes fatores da decomposição PARAFAC geralmente é realizada por minimização da seguinte função não-linear quadrática:

$$f(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \left\| \left\| \mathcal{X} - \sum_{q=1}^Q \mathbf{A}_{.q} \circ \mathbf{B}_{.q} \circ \mathbf{C}_{.q} \right\|_F \right\|^2. \quad (4.33)$$

O algoritmo dos Mínimos Quadrados Alternados (ALS) é uma solução clássica de minimizar esta função de perda. É um algoritmo iterativo que alterna entre as estimações de  $\mathbf{A}$ ,  $\mathbf{B}$  e  $\mathbf{C}$ . Em outras palavras, o algoritmo ALS converte um problema não-linear (na verdade, trilinear) de otimização de três independências lineares em mínimos quadrados (LS).

Cada interação é composta de três etapas de estimação dos Mínimos Quadrados. Em cada etapa, uma matriz fator, declara,  $\mathbf{A}$  é atualizado enquanto os outros dois ( $\mathbf{B}$  e  $\mathbf{C}$ ) são fixados em seus valores nas etapas anteriores de estimação. O algoritmo

faz uso da fatorização de Khatri-Rao das matrizes desdobradas  $\mathbf{X}_1$ ,  $\mathbf{X}_2$  e  $\mathbf{X}_3$  dado em (1.24). Podem ser resumidas da seguinte forma:

i. Fixo  $i = 0$ ;

Inicialize randomicamente  $\hat{\mathbf{B}}_{(i=0)}$  e  $\hat{\mathbf{C}}_{(i=0)}$ ;

ii.  $i = i + 1$ ;

iii. A partir de  $\mathbf{X}_3$  e usando  $\hat{\mathbf{B}}_{(i-1)}$  e  $\hat{\mathbf{C}}_{(i-1)}$  encontrando uma estimativa de  $\mathbf{A}$  resolvendo o problema dos mínimos quadrados seguintes:

$$\hat{\mathbf{A}}_{(i)} = \underset{\mathbf{A}}{\operatorname{argmin}} \left\| \mathbf{X}_3 - (\mathbf{B}_{i-1} \diamond \mathbf{C}_{(i-1)}) \mathbf{A}^T \right\|_F^2$$

a solução é dada por:  $\hat{\mathbf{A}}_{(i)}^T = (\hat{\mathbf{B}}_{(i-1)} \diamond \hat{\mathbf{C}}_{(i-1)})^\dagger \mathbf{X}_3$

iv. A partir  $\mathbf{X}_1$  e usando  $\hat{\mathbf{A}}_{(1)}$  e  $\hat{\mathbf{C}}_{(1)}$ , encontrar uma estimativa de  $\mathbf{B}$  resolvendo o problema dos mínimos quadrados seguintes:

$$\hat{\mathbf{B}}_{(i)} = \underset{\mathbf{B}}{\operatorname{argmin}} \left\| \mathbf{X}_1 - (\mathbf{C}_{i-1} \diamond \mathbf{A}_{(i-1)}) \mathbf{B}^T \right\|_F^2$$

a solução é dada por:  $\hat{\mathbf{B}}_{(i)}^T = (\hat{\mathbf{C}}_{(i-1)} \diamond \hat{\mathbf{A}}_{(i)})^\dagger \mathbf{X}_1$

v. A partir  $\mathbf{X}_2$  e usando  $\hat{\mathbf{A}}_{(1)}$  e  $\hat{\mathbf{B}}_{(1)}$ , encontrar uma estimativa de  $\mathbf{C}$  resolvendo o problema dos mínimos quadrados seguintes:

$$\hat{\mathbf{C}}_{(i)} = \underset{\mathbf{C}}{\operatorname{argmin}} \left\| \mathbf{X}_2 - (\mathbf{A}_{i-1} \diamond \mathbf{B}_{(i-1)}) \mathbf{C}^T \right\|_F^2$$

a solução é dada por:  $\hat{\mathbf{C}}_{(i)}^T = (\hat{\mathbf{A}}_{(i)} \diamond \hat{\mathbf{B}}_{(i)})^\dagger \mathbf{X}_2$ ;

vi. Repita os passos 2-5 até a convergência.

A convergência é declarada na  $i$ -ésima interação quando o erro entre o verdadeiro tensor e a versão construída, a partir da matriz fator estimada, não se altera significativamente entre as interações  $i$  e  $i + 1$ . Uma medida erro na  $i$ -ésima interação poder ser calculada a partir da seguinte fórmula:

$$e_{(i)} = \left\| \mathbf{X}_1 - (\mathbf{C}_{(i)} \diamond \mathbf{A}_{(i)} \mathbf{B}_{(i)}^T) \right\|_F$$

A convergência na  $i$ -interação pode ser declarada quando  $\| e_{(i+1)} - e_{(i)} \| < \delta$ , onde  $\delta$  é um valor limite prescrito (por exemplo,  $\delta = 10^{-6}$ ).

Uma atualização condicional de qualquer matriz dada pode melhorar ou manter, mas não pode piorar o ajuste atual. O algoritmo sempre monotonicamente converge (pelo menos) para um mínimo local. Contudo, o algoritmo ALS é fortemente dependente na inicialização e da convergência de um mínimo global algumas vezes pode ser lento. Além disso, a convergência do algoritmo pode, em alguns casos, cair em regiões de "pântano", durante o qual a velocidade da convergência é muito pequena e o erro entre as duas consecutivas interações não diminui. No caso de evitar uma interrupção prematura do algoritmo, uma prática comum é impor um valor com o mínimo de aceitação  $e_{(i)}$ , para o qual a convergência global não será assumida se ainda não for alcançado (mesmo quando  $\| e_{(i+1)} - e_{(i)} \| < \delta$ ).

### Métodos e Algoritmos Alternativos

Existem inúmeros algoritmos do ALS proposto na literatura. (Ver demais trabalhos para uma revisão crítica em alguns algoritmos alternativos). A fim de aliviar os problemas de convergência lenta causada por uma inicialização aleatória do algoritmo, uma análise da própria solução pode ser usada. Esta solução também é conhecida como decomposição trilinear direta. Consiste na obtenção de uma primeira estimação das matrizes-fatores da decomposição através da construção de um problema de autovalor generalizado (ou problema de diagonalização) a partir de duas fatias do tensor. A inicialização baseada na própria solução, além de ser limitada a tensores com apenas duas fatias em um dos modos, requer que ambas as matrizes fatores sejam rank-coluna completo e o terceiro não contenha elementos igual a zero. O trabalho propõe uma generalização da solução própria análise ao tensor com mais do que duas fatias, ligando a estimativa das matrizes fator do PARAFAC ao problema da diagonalização da matriz simultânea.

Outra forma de melhorar a velocidade do algoritmo ALS é baseado do método de compressão Tucker-3. Este método é útil quando a dimensão do tensor é grande. Foi proposto na literatura um algoritmo para acelerar a convergência do ALS. Este algoritmo aplica-se ao método de compressão Tucker-3 seguido por uma inicialização baseada na própria análise. A convergência do algoritmo do ALS pode também ser melhorada por meio do chamado de método Pesquisa Linha Melhorada (ELS). O método ELS será mostrado para ser útil quando a decomposição do tensor é afetada pelo fator degeneração. O método ELS foi também usado para o fator de estimação

do bloco de decomposição do tensor. Um método Levenberg-Marquadt é proposto.

### Conexão entre o PARAFAC e Tucker-3

Embora Tucker-3 e o PARAFAC sejam conceitualmente decomposições tensoriais diferentes, é possível aproximá-los em alguns casos. Em outras palavras, o PARAFAC pode ser interpretado como um caso especial do Tucker-3 bem como o Tucker-3 poder ser visto como uma versão *restringida* do PARAFAC. Isto indica que ambas decomposições poderia, em princípio, ser usadas para modelar os dados do mesmo tensor de dados, embora os cálculos de cada um sejam diferentes.

A equivalência conceitual entre o Tucker-3 e o PARAFAC foram bem discutidas por Harshman. Ele também apontou que não só apenas o modelo PARAFAC poderia ser equivalentemente, mas também afirmou ser um caso especial do Tucker-3 e vice-versa, mas ambos podem ser combinados em diferentes formas, resultando em uma decomposição híbrida que simultaneamente desfruta propriedades do Tucker-3 e do PARAFAC. Esta combinação pode ser útil em casos onde se está interessado na caracterização/modelagem de tensores com a mais complexa estrutura inerente.

Voltando ao modelo Tucker-3 (1.7) e considerando que  $P = Q = R$ , nós temos:

$$x_{i_1, i_2, i_3} = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R a_{i_1, p} b_{i_2, q} c_{i_3, r} g_{p, q, r} = \sum_{q=1}^Q a_{i_1, p} b_{i_2, p} c_{i_3, q} g_{q, q, q} \quad (4.34)$$

Note que vamos cair no modelo PARAFAC padrão quando  $g_{q, q, q} = \delta_{q, q, q}$ . Em outras palavras, o PARAFAC de terceira ordem é uma caso especial do Tucker-3 com "superdiagonal", do tensor núcleo, sendo assim,  $\mathcal{G} = \mathcal{I}_Q$ , no qual todos os seus elementos são zeros em todas as posições exceto na superdiagonal principal, em que todos os elementos são um.  $\mathcal{I}_Q$  é chamado de *Tensor Identidade*. Note que os elementos da superdiagonal de  $g_{q, q, q}$  não precisam ser restritos a um, desde sempre se pode absorver  $g_{q, q, q}$  em qualquer das três matrizes fator e ainda ter uma representação equivalente ao PARAFAC como um caso especial. Unicidade neste caso especial é possível, desde que as condições necessárias e suficientes do PARAFAC sejam satisfeitas.

Agora, vamos entrar na direção oposta e mostrar agora que a decomposição Tucker-3 pode ser interpretada como um PARAFAC restringido. Reescrito (1.7) na

sua forma equivalente (1.22):

$$x_{i_2, i_2, i_3} = \sum_{p=1}^P \sum_{q=1}^Q a_{i_1, p} b_{i_2, q} h_{p, q, i_3}, \quad (4.35)$$

com  $h_{p, q, i_3} = [\mathcal{G} \bullet_3 \mathbf{C}]_{p, q, i_3}$ . Define  $\tilde{\mathbf{C}} \in \mathbb{C}^{I_3 \times PQ}$  como uma matriz desdobrada do tensor núcleo  $h_{p, q, i_3}$  da seguinte maneira:

$$\tilde{c}_{i_3, p, q} = [\tilde{\mathbf{C}}]_{i_3, (q-1)P+p}. \quad (4.36)$$

Defini também a matrizes fator *aumentada*  $\tilde{\mathbf{A}} \in \mathbb{C}^{I_1 \times PQ}$  e  $\tilde{\mathbf{B}} \in \mathbb{C}^{I_2 \times PQ}$  como:

$$\tilde{\mathbf{A}} = \underbrace{[\mathbf{A}, \dots, \mathbf{A}]}_{Q\text{-termos}}, \quad \tilde{\mathbf{B}} = \underbrace{[\mathbf{B}_{\cdot 1}, \dots, \mathbf{B}_{\cdot 1}]}_{P\text{-termos}}, \dots, \underbrace{[\mathbf{B}_{\cdot Q}, \dots, \mathbf{B}_{\cdot Q}]}_{P\text{-termos}} \quad (4.37)$$

onde  $\tilde{\mathbf{A}}$  é uma concatenação de  $Q$  matrizes idênticas  $\mathbf{A}$ , enquanto  $\tilde{\mathbf{B}}$  é uma matriz onde cada coluna de  $\mathbf{B}$ , é denotada por  $\mathbf{B}_{\cdot q}$ ,  $q = 1, \dots, Q$  é repetido  $P$  vezes. Tomando esta definição em consideração, uma decomposição Tucker-3 pode ser representada pela seguinte decomposição PARAFAC de terceira ordem de fatores  $F = PQ$ :

$$x_{i_1, i_2, i_3} = \sum_{f=1}^F \tilde{a}_{i_1, f} \tilde{b}_{i_2, f} \tilde{c}_{i_3, f}, \quad (4.38)$$

onde os fatores do primeiro e do segundo modo  $\tilde{a}_{i_1, f} = [\tilde{\mathbf{A}}]_{i_1, f}$  e  $\tilde{b}_{i_2, f} = [\tilde{\mathbf{B}}]_{i_2, f}$  são restringidos na seguinte forma:

$$[\tilde{\mathbf{A}}]_{i_1, (q-1)P+p} = [\mathbf{A}]_{i_1, p}, \quad [\tilde{\mathbf{B}}]_{i_2, (q-1)P+p} = [\mathbf{B}]_{i_2, q}, \quad (4.39)$$

onde  $f = (q-1)P+p$ ,  $p = 1, \dots, P$ ,  $q = 1, \dots, Q$ . Assim, a decomposição Tucker-3 sempre pode ser incorporado em um PARAFAC, onde o primeiro e o segundo modo das matrizes fatores têm repetidos vetores de carga enquanto o terceiro modo da matriz fator é uma matriz de desdobramento do núcleo do Tucker-3. Os padrões de repetições em  $\tilde{\mathbf{A}}$  e  $\tilde{\mathbf{B}}$  são dados por (1.40). Eles produzem todos as combinações possíveis de fatores através do produto externo dos fatores necessárias para caracterizar completamente a estrutura de interação do Tucker-3.

A decomposição PARAFAC equivalente (1.39) não é única, em geral, uma vez que a singularidade não admite um par de colunas proporcionais em qualquer das matrizes fatores. Ambas  $\tilde{\mathbf{A}}$  e  $\tilde{\mathbf{B}}$  são tais que  $k_{\tilde{\mathbf{A}}} = k_{\tilde{\mathbf{B}}} = 1$  e a condição necessárias (1.30) não comprova a unicidade. Portanto, qualquer representação do PARAFAC de uma estrutura geral Tucker-3 irá mostrar as mesmas indeterminações que a representação do Tucker-3.

# Capítulo 5

## Procedimentos Metodológicos

### 5.1 Caracterização dos Dados

---

Para o desenvolvimento deste trabalho foi criada uma base de dados com valores randômicos através do software MatLab tendo a estrutura composta por supostamente 200 alunos do Curso de Licenciatura Plena em Física na modalidade presencial da Universidade Federal do Ceará (UFC). Estes supostos alunos responderam a um questionário no qual deram notas, baseadas na escala Likert, em relação às características do curso e do professor, fundamentadas no SEEQ. Todos os alunos estão cursando ou já cursaram as disciplinas de Física 1, Métodos de Física 1, Cálculo 1, e Estudo Sócio-político da Educação, todas as disciplinas citadas são ofertadas exclusivamente no primeiro semestre, porém as disciplinas de Cálculo 1 e Estudo Sócio-político são ofertadas de forma semestral tendo em vista que essas disciplinas são ofertadas para todos os cursos do Centro de Ciências e para os cursos de Licenciatura, respectivamente.

#### 5.1.1 Graduações presenciais de Licenciatura Plena

A carga horária do curso de Licenciatura em Física possui um total de 2800 horas, a serem integralizadas num prazo de 4 anos, sendo que desse total de horas são divididos em: Disciplinas obrigatórias: 1600 horas, disciplinas optativas: 200 horas, disciplinas de livre escolha: até 360 h (20% do total de 1800h), 400 h de prática como componente curricular (Métodos de Ensino), 400h de prática de Estágio Curricular Supervisionado e 200h de atividades complementares.

### 5.1.2 Conteúdos das Atividades Analisadas

Os conteúdos das supostas disciplinas em análise, que serão apresentados a seguir:

#### i. Física Fundamental I: Possui a carga horária de 128 horas/semestre

**Tabela 5.1:** O conteúdo visto na parte teórica da disciplina

1	Medições
2	Vetores
3	Movimentos Retilíneo
4	Movimento em duas e três dimensões
5	Força e Movimento
6	Trabalho e Energia Cinética
7	Conservação da Energia
8	Sistemas de Partículas

**Tabela 5.2:** Parte experimental é composta pelo conjunto de seis experimentos

1	Paquímetro
2	Micrômetro
3	Movimento Retilíneo Uniforme
4	Movimento Retilíneo Uniformemente Variado
5	Lei de Hooke e Associação de Molas
6	2ª Lei de Newton

#### ii. Cálculo Diferencial Integral I: A Tabela 5.3 apresenta a ementa da disciplina, possui 96 horas/semestre

**Tabela 5.3:** Conteúdos da disciplina de Cálculo Diferencial Integral I.

1	Funções reais de uma variável real
2	Limite e continuidade
3	Funções trigonométricas, exponencial e logarítmica
4	Derivadas e suas aplicações
5	Primitivas



**iii. Método de Ensino Física I:** Possui a carga horária de 32 h/semestre.

Seleção e avaliação de metodologias, estratégias e recursos adequados ao ensino, nas escolas de ensino fundamental e médio, dos conteúdos de mecânica desenvolvidos em Física Fundamental I. Identificação de conteúdos e objetivos, estabelecendo suas implicações na educação científica e no desenvolvimento curricular. Proposição e desenvolvimento de estratégias, materiais e instrumentos de avaliação. Análise dos livros didáticos adotados nas escolas de ensino fundamental e médio.

**iv. Estudo Sócio-Político da Educação**

Descrição Nesta disciplina teremos oportunidade de refletir sobre a educação e suas determinações e influências do contexto social e histórico e inter-relações com a sociedade. Em sua ementa essa disciplina se propõe ao estudo da contribuição das ciências sociais e humanas para a compreensão do fenômeno educativo e sua aplicação no processo de formação do educador. As disciplinas Sociologia e História dispõem de um conjunto de proposições desses campos do saber e podem fornecer os fundamentos ou os princípios básicos dessas áreas de conhecimentos no sentido de fornecer o apoio, ou os subsídios para se entender a Educação como um fenômeno social e histórico e, com isso, pode contribuir para que o educador compreenda sua ação educativa como um processo que sofre.

**Objetivos:**

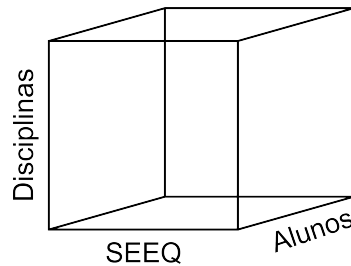
- (a) Propiciar ao estudante do Curso de Física o acesso e a aquisição de conhecimentos sobre os fundamentos sociais e históricos da educação;
- (b) Analisar a educação como um fenômeno que ocorre no tempo, no espaço e na rede complexa das relações sociais que tecem a história das sociedades humanas;
- (c) Conhecer as dimensões sociais e históricas do processo educativo no Brasil.

**5.2 Estrutura dos Dados em Três Dimensões**

---

Os dados foram arranjados de tal modo que criasse uma matriz multidimensional, também chamado de tensor, onde as disciplinas estão sendo representadas nas linhas

da matriz, as características do SEEQ estão nas colunas e as suspostas respostas dos alunos estão nas fatias do tensor, frisando que os dados foram criados através do programa MatLab usando número randômicos. A matriz tem dimensão  $4 \times 10 \times 200$ , 4 disciplinas, 11 características do SEEQ e avaliação de 200 alunos, para facilitar a representação pictórica iremos chama a dimensão das discilpinas de  $I$ , a dimensão das características do SEEQ de  $J$  e  $K$  avaliação do aluno. Figura 5.1.



**Figura 5.1:** Representação pictórica dos dados das disciplinas na forma matricial.

Escala Likert é uma escala de resposta psicométrica usada habitualmente em questionários, e é a escala mais usada em pesquisas de opinião. A escala é bipolar, medindo ou uma resposta positiva ou negativa a uma afirmação. Às vezes são usados quatro itens, o que força o sujeito pesquisado a uma escolha positiva ou negativa, uma vez que a opção central "Indiferente" não existe. Escalas de Likert podem estar sujeitas a distorções por diversas causas. Sujeitos perguntados podem evitar o uso de respostas extremas, concordar com afirmações apresentadas ou tentar mostrar a si ou a suas organizações de um modo mais favorável. O desenho da escala com respostas mais balanceadas pode resolver a questão dos desvios por aceitação às afirmações. O formato dos *itens* da escala Likert:

- i. Muito Ruim
- ii. Ruim
- iii. Moderado
- iv. Bom
- v. Muito Bom

Dessa maneira as fatias frontais representam as respostas do questionário dos alunos com relação a cada disciplina, ou seja, a primeira fatia representa a opinião

de um suposto aluno com relação as quatro disciplinas em questão, no caso serão Física Fundamental 1, Cálculo Diferencial Integral I, Método de Ensino de Física 1 e Estudo Sócio-político da Educação e dando a nota, baseada na escala Likert. Com relação às características do SETE foram divididas em categorias:

- i. Aprendizagem. Se o estudante está intelectualmente desafiante e estimulante; Se o interesse do aluno tem aumentado com uma consequência desse curso
- ii. Enstusiasmo. O professor está entusiamado com o curso
- iii. Organização. As explicações eram bem claras; O material disponibilizado é de fácil entedimento.
- iv. Interação do grupo. Os alunos foram incentivados a participar das discussões em aula.
- v. Afinidade Individual. O professor trata os alunos de forma amigável.
- vi. Abrangência. Professor contrastava as implicações de várias teorias.
- vii. Exames. Comentários sobre exames / materiais classificados foi de grande valia
- viii. Trabalhos Leituras Obrigatórias e textos foram importantes.
- ix. Geral Em comparação com outros cursos que você tomou conhecimento na UFC
- x. Geral Em comparação com outros professores que você teve na UFC

Os alunos avaliaram estas categorias respondendo questões que estão relacionadas com as categorias, atribuindo notas 1 a 5, baseadas na escala likert, onde cada número representa, respectivamente, muito ruim, ruim, moderado, bom, muito bom. As notas foram distribuídas seguindo o conhecimento empírico educacional. As disciplinas de Física Introdutória 1 e Cálculo Diferencial Integral 1 apresentaram a mesma moda 4, pois apresentam as mesmas características educacionais, dificuldade, índice de aprovação. A disciplina de Métodos de Ensino de Física 1 avalia metodologias, estratégias e recursos adequados dos conteúdos desenvolvidos em relação a Física Introdutória 1, e estando relacionadas. Porém a disciplina de

Estudo Sócio-Político da Educação (ESPE), praticamente não existe relação com as disciplinas anteriores, pois é ela é o ofertada pelo departamento de Pedagogia.

A disciplina de Métodos de Ensino de Física 1 avalia metodologias, estratégias e recursos adequados dos conteúdos desenvolvidos de Física Introdutória 1.

### 5.3 Estrutura dos Dados em Duas Dimensões

---

Devido o modo de como os dados estão arranjados foi necessário fazer a médias da dimensão dos aluno para aplicação do PCA juntar á decomposição PARAFAC para verificar qual a métodos é melhor para esta situação.

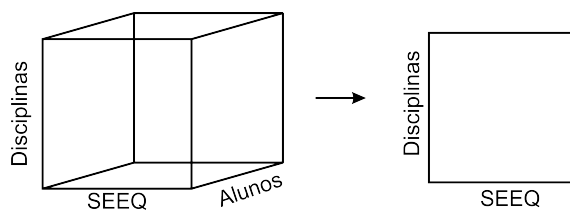


Figura 5.2: Representação pictórica da redução de dimensão.

### 5.4 Criação e Tratamento dos Dados

---

Como foi dito anteriormente os dados utilizados no desenvolvimento deste trabalho foram criados através do software MatLab, porém como os dados ficariam muito dispersos logo dificultaria muito a suposta interpretação, então foi criado um intervalo de variação das notas para cada disciplina, ou seja, as notas foram viciadas de tal maneira que induzisse a uma interpretação esperada, já que o trabalho tenho por objetivo apenas demonstrar a utilização da ferramenta de avaliação SETE.

As disciplinas de Física Fundamental 1 e Cálculo Diferencial Integral I foram induzidas a variarem no intervalo de números inteiros de 1 a 4, já que apresentam características semelhantes com relação as dificuldades encontradas pelos alunos durante o decorrer da disciplina. Métodos de Física 1 variou no intervalo de números inteiros 3 a 5, pois é uma disciplina que apenas tem o alvo somente orientar sobre a metodologia usada em sala de aula, entretanto como a disciplina não tem um cronograma a ser seguido, fica a cargo do professor qual a melhor forma de conduzi-la, podendo o aluno enfrentar dificuldades. Por outro lado, a disciplina de Estudo Sócio Político da Educação é uma disciplina que apenas envolve muita leitura e compreensão textual.

Outra consideração importante foi realizada em relação ao questionário, para cada característica do SETE existem perguntas que devem ser respondidas de acordo com a escala Likert, logo faz-se a média das perguntas, correspondente a cada característica, para obter a pontuação da características através das tais perguntas.

## 5.5 Descrição do Processo de Modelagem no MATLAB

---

### 5.5.1 Criando o tensor

Obtem-se as matrizes correspondente às respostas dos alunos, às fatias do tensor, usando a seguinte fórmula:

$$\text{round}((b - a) * \text{rand}(c, d) + a);$$

onde

- **round**: Indica que o intervalo será de números inteiros;
- (a): O valor mínimo do intervalo;
- (b): O valor máximo do intervalo;
- rand**: Criação de números de forma randômica;
- (c,d): A dimensão da matriz onde *c* e *d* representam as linhas e colunas, respectivamente;

### 5.5.2 Descrição da Modelagem do PCA

Para a realização da decomposição PCA foi utilizado o *software* MATLAB. Essa ferramenta disponibiliza a comando  $[COEFF, SCORE, latent, tsquare] = \text{princomp}(X)$  que possui alguns parâmetros de entrada e saída. Como será apresentada a seguir na forma de código MATLAB.

#### i. Parâmetros de Entrada

**X**: Representa os dados.

**COEFF**: Matriz de Coeficientes, também conhecido como "loadings".

**SCORE**: Representa X no espaço da componentes principais.

**latent**: É o vetor que contém os autovalores da matriz de covariância de X

**tsquare:** É uma medida das distância multivariadas de cada observação a partir do centro do conjunto de dados.

### 5.5.3 Descrição da Modelagem no *N-way Toolbox*

Para a realização da decomposição PARAFAC foi utilizado o *software* MATLAB juntamente com o *N-way Toolbox*. Essa ferramenta disponibiliza a função *parafac.m* que possui alguns parâmetros de entrada e saída. Como será apresentada a seguir na forma de código MATLAB.

```
[Factors, it, err, corcondia] = parafac(X, Fac, Options, const, OldLoad, FixMode, Weights)
```

Algumas restrições podem ser utilizadas em modos individuais para a obtenção de soluções não-negativas, ortogonais, ou unimodais. Nessa função todos os vetores de carga, exceto no primeiro modo, são normalizados, de modo que toda a variância é mantida no primeiro modo. Ao término das interações o primeiro componente é considerado o mais importante. Para a utilização desse algoritmo consideramos algumas restrições para que ocorresse uma melhora no desempenho, o modelo final utilizado pode ser visto no Apêndice B. A seguir serão apresentados os parâmetros de entrada e saída da função *parafac.m*:

#### i. Parâmetros de Entrada

**X:** Conjunto de dados multidimensional.

**Fac:** Número de componentes.

**Options:**

- 1 - Critério de convergência;
- 2 - Método de inicialização do algoritmo;
- 3 - Opções de plotagem das componentes;
- 4 - Não utilizada pelo usuário;
- 5 - Mostra o ajuste entre os dados reais e os estimados;
- 6 - Número máximo de iterações.

**const:** Colocar restrições nos modos das matrizes.

- 0 - Sem restrição;
- 1 - Ortogonalidade;
- 2 - Não-negatividade;
- 3 - Unimodal (e não-negatividade);

**OldLoad:** Se as estimativas iniciais das cargas estiverem disponíveis.

**FixMode:** Vetor binário de mesma dimensão de  $X$  (matriciado modo-1) que estabelece quais são as componentes fixas no modelo.

**Weights:** Matriz que relaciona os pesos entre as componentes.

## ii. Parâmetros de Saída

**Factors:** Matrizes de carregamento.

**it:** Número de interações.

**err:** Erro atingido pela convergência do algoritmo.

**corcondia:** Teste aplicado para testar a aplicabilidade do modelo PARAFAC para o conjunto de dados em questão.

# Capítulo 6

## Análise e Discussão dos Resultados

### 6.1 Análise dos Dados

---

A análise das componentes principais e também multidimensional, que foram modeladas através do software MATLAB a partir da base de dados criados (4 disciplinas, 10 características do SEEQ e 200 alunos) cujos alunos participaram das disciplinas Física Introdutória I, Cálculo Diferencial e Integral I, Métodos de Ensino de Física 1 e Estudo Sócio Político da Educação na modalidade presencial da Universidade Federal do Ceará (UFC). Através dessa base de dados foi realizada uma análise do desempenho dos alunos em todas as disciplinas e atividades, os resultados obtidos serão discutidos a seguir.

#### 6.1.1 Análise das Componentes Principais

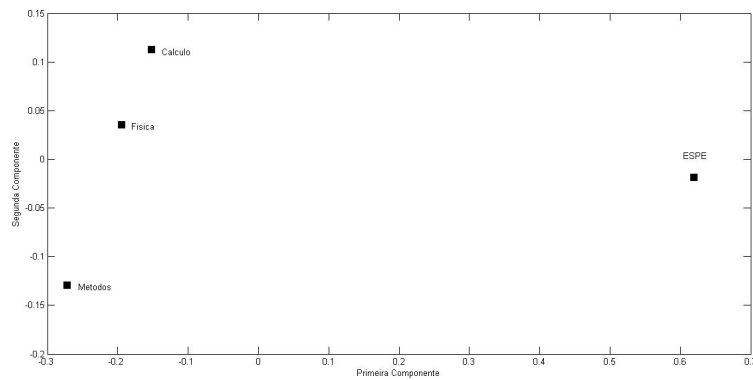
A utilização do PCA gera duas matrizes chamadas de matrizes de coeficientes e matrizes dos *scores*, que *scores* são autovetores, normalizados e ortogonais e formam o espaço do PCA. Através da combinação linear dos autovetores,

$$Y_j = e'_j X = e_{j1}X_1 + e_{j2}X_2 + \dots + e_{jp}X_p$$

, para que tenhamos as componentes principais.

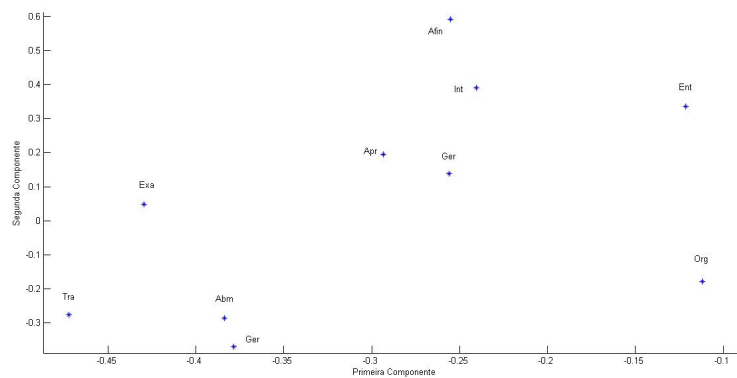
Ao analisarmos a matriz de coeficientes ( $4 \times 10$ ), observamos que ela nos fornece informações sobre o comportamento das disciplinas. Existe um agrupamento próximo à segunda componente o que nos leva a seguinte interpretação: o grupo é composto pelas disciplinas que há relação curricular é o caso de Cálculo Diferencial





**Figura 6.1:** Matriz de coeficientes.

e Integral 1 e Física Fundamental 1, entretanto a Métodos de Ensino de Física 1 está distante em relação as outras duas, mas tem características semelhante, citadas anteriormente, por isso está ligada a segunda componente



**Figura 6.2:** Matriz de scores.

Apr: Aprendizagem; Ent: Entusiasmo; Org: Organização; Int: Interação de Grupo; Afin: Afinidade Individual; Abrn: Abrangência; Trab: Trabalho; Ger: Geral; Ger: Geral.

No gráfico dos escores podemos ver que existem três grupos: Exames, trabalhos, abrangência e geral estão ligadas pelo fato de as duas primeiras estarem relacionadas com as atividades do curso estando a abrangência às anteriores porquanto se refere as explicações bem claras ajudando o aluno na resolução de exercício.

O segundo grupo possui: Afinidade Individual, Aprendizagem, Interação do Grupo Trabalhando em equipe, o estudante exercita uma série de habilidades. Ao mesmo tempo em que estuda o conteúdo das disciplinas, ele aprende a escolher, a

avaliar e a decidir. "Nesse tipo de tarefa, treina-se a capacidade de ouvir e respeitar opiniões diferentes", explica Stella Galli. A lista ainda inclui saber argumentar e dividir tarefas - competências essenciais para a vida adulta. "Ao crescer, o jovem vai ter de conviver com pessoas diferentes dele", explica Adilson Garcia, diretor do Colégio Vértice, de São Paulo. "O preparo para vida depende de saber se comunicar, ouvir, respeitar interlocutores e isso só se aprende fazendo"

O terceiro Grupo: Organização e Entusiasmo. O entusiasmo pela educação teve caráter quantitativo, visando a expansão da rede escolar e analfabetizar o povo. O otimismo pedagógico teve caráter qualitativo, importava-se com a otimização do ensino, ou seja, com a melhoria das condições didáticas e pedagógicas da rede escolar (Ghiraldelli Jr., 2003).

### 6.1.2 Análise Multidimensional

Seja  $\mathcal{X}$  um tensor tridimensional de tamanho  $I \times J \times K$ , ( $4 \times 10 \times 200$ ). Consideramos a decomposição PARAFAC, cuja ideia é fatorar o tensor  $\mathcal{X}$  em três matrizes  $\mathbf{A} \in \mathbb{R}^{I \times R}$ ,  $\mathbf{B} \in \mathbb{R}^{J \times R}$  e  $\mathbf{C} \in \mathbb{R}^{K \times R}$

$$x_{ijk} = \sum_{r=1}^R a_{ir} b_{jr} c_{kr} + e_{ijk}$$

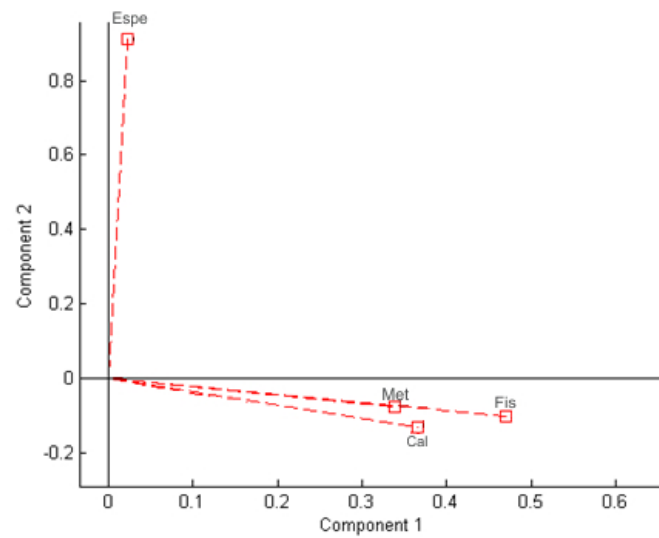
onde  $\mathcal{E} \in \mathbb{R}^{I \times J \times K}$  é o tensor residual.

Um método de encontrar tais matrizes  $\mathbf{A}$ ,  $\mathbf{B}$  e  $\mathbf{C}$  é minimizar a seguinte função custo:

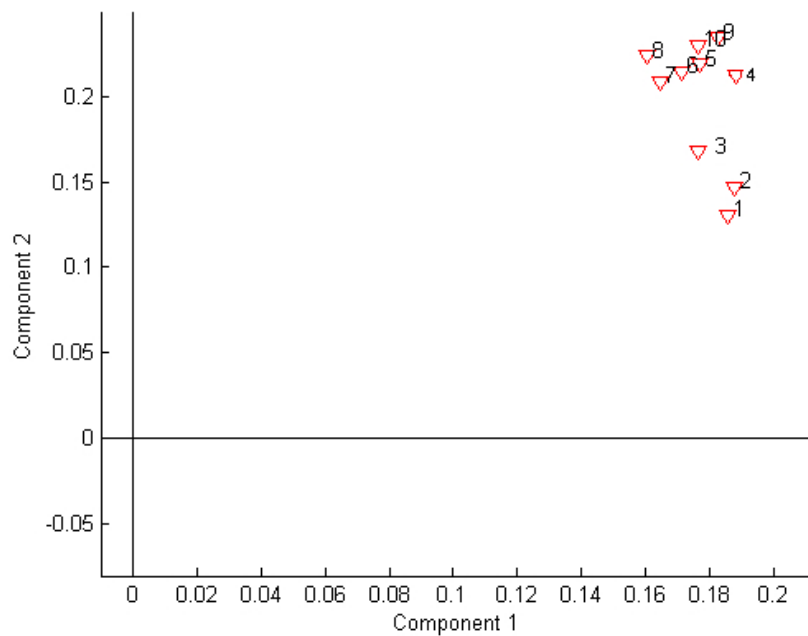
$$f(A, B, C) = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K \left( x_{ijk} - \sum_{r=1}^R a_{ir} b_{jr} c_{kr} \right) \quad (6.1)$$

Na figura 6.3 pode-se perceber que a disciplina de Estudo Sócio-Político da Educação (ESPE) mostrou-se praticamente perpendicular com as demais, o que foi induzido. O fato de escolhermos esta disciplina para ficar distante das demais foi a sua disparidade com a demais, em relação a componente curricular, já que as três têm uma correlação, Física e Cálculo estão diretamente ligadas, já a disciplina de métodos está ligada com Física, pois o seu objetivo é: avaliar a metodologia, estratégias e recursos adequados no ensino médio do conteúdo, neste de Física 1.

Os numerais correspondem as seguintes características, respectivamente: Apr:



**Figura 6.3:** Matriz A da decomposição do PARAFAC.

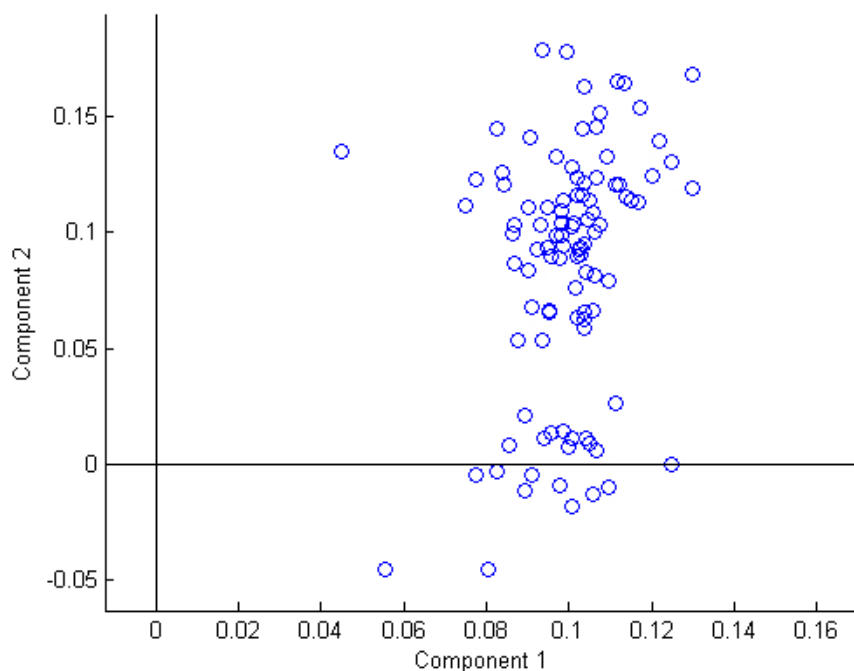


**Figura 6.4:** Matriz B da decomposição do PARAFAC.

Aprendizagem; Ent: Entusiasmo; Org: Organização; Int: Interação de Grupo; Afin: Afinidade Individual; Abrn: Abrangência; Trab: Trabalho; Ger: Geral; Ger: Geral

As 11 questões podem ser analisadas através da matriz  $\mathbf{B} \in \mathbb{R}^{I \times R}$  gerada pelo PARAFAC, onde, no nosso caso,  $J=11$  e  $R=2$ . A figura 6.4 mostra a plotagem das supostas características em função das 2 componentes geradas pelo PARAFAC.

Observa-se uma maior variabilidade em relação à segunda componente do que na primeira. Outra observação que as questões tendem a estar acumuladas numa região, revelando um grau de proximidades entre elas. Lembrando que esses dados são de uma simulação controlada. As questões na ordem de suas enumerações, observaremos as questões na ordem de suas enumerações, observaremos uma proximidade entre algumas questões, o que o questionário SETE indica a ser feito para que haja correlação entre as questões, abordando assuntos parecidos entre algumas delas.



**Figura 6.5:** Matriz  $C$  da decomposição do PARAFAC.

A figura 6.5 representa  $C \in \mathbb{R}^{K \times R}$  mostra a distribuição dos alunos diante das duas componentes geradas pela decomposição PARAFAC. Observa-se a existência de dois agrupamentos, o que foi induzido, o primeiro agrupamento superior mostra uma quantidade de alunos. A segunda, pode se dizer em dados reais, que o agrupamento superior são de alunos que cursaram todas as disciplinas e o agrupamento inferior corresponde aos demais.

## Considerações Finais

Neste trabalho foi realizada uma revisão teórica a respeito da Avaliação Multidimensional utilizando as seguintes teorias: (MARSH; DUNKIN, 1991), e de (MURRAY H. G., 1990) onde se avalia a estabilidade e processos de padronização do SETE (STEVENS, 1986) e por Tabachnick e Fidell, para uma melhor interpretação. Foi feita uma revisão bibliográfica sobre álgebra multilinear, apresentando algumas definições, propriedades, notações e operações básicas utilizadas nessa área. Por fim, apresentamos algumas ferramentas computacionais importantes no desenvolvimento de trabalhos baseados na álgebra multilinear, dentre elas o *N-way Toolbox*.

Avaliação multidimensional mostrou ser importante para o desenvolvimento de uma instituição de ensino e melhoria no sistema educacional, seja ele na instituição pública ou privada. Deve ser visto como um método de identificação de problemas de forma direta, pois é através do questionário que os problemas são suscitados. A partir do estudo teórico feito neste trabalho percebemos que sua aplicação traz melhorias em todas as esferas. Estudantis, administrativas, educacionais e governamentais, garantido o anonimato do discente, deixando-o livre de qualquer constrangimento, além do mais o professor passa a ser avaliado de forma crítica e podendo auxiliá-lo para promoção profissional dentro de sua instituição.

Foi mostrado o método de redução de dimensão de componentes chamado de Análise das Componentes Principais o qual visa resumir a informação em um número pequeno de componentes e verificar a existência de grupo, como conduzido neste trabalho, mostrando ser uma análise que trabalha com mais de uma variável dependente além de ser uma análise exploratória e útil em estudos de comunidade e o

uso de variáveis latente, porém, neste caso, há um perda de dimensionalidade tendo em vista que Análise das Componentes Principais é uma análise multidimensional, sendo assim foi necessário fazer um ajuste na dimensão dos alunos. Indica-se futuramente que se use Análise Fatorial para a confirmação de tais grupos citados neste trabalho.

Para comparar com outro tipo de análise foi utilizada uma decomposição que visa a redução de dimensionalidade com a menor perda possível através da equação 6.1 e uma das propriedades mais interessantes do PARAFAC é a unicidade nas componentes, contrariamente na decomposições bilinear (matriz), que em geral não são únicos para ranks superiores a um, (matrizes de rank-um até atingir um fator de escala), facilitando a interpretação dos dados e a utilização de um método de validação chamado de Corcôndia, não citado neste trabalho, no qual visa se o número de fatores estipuladas pelo usuário é possível com aqueles dados, tornando a decomposição mais estável e mais confiável e a consideração da dimensão alunos que não foi possível visualizar, pois o PCA é bidimensional. E para trabalhos futuros a utilização de outras decomposições.

Por fim, o ensaio mostrado neste trabalho teve o intuito de apresentar um método de avaliação multidimensional e compará-lo com diferentes tipos de análises, PCA e PARAFAC. Além disso mostram a importância da utilização, apesar de ter sido necessário fazer ajuste nos dados para utilizar o PCA. Os dados foram feitos com dados matemáticos, por não ter sido viabilizado aplicação de dados reais, porém os próximos trabalhos que virão deverão utiliza-los, porém a modelagem computacional poderá ser empregada em pesquisas, se o objetivo for apenas um teste de uma ferramenta visto que mostrou-se controle sobre os dados.

# Apêndice **A**

## Código MATLAB

### A.0.3 Código Utilizado para Criação das Matrizes

```
clc;
clear all;
%% Fisica Fundamental:
a=round((5-1)*rand(10,200)+1);
for cont=0:1400;
for j=round((200-1)*rand(1,1)+1);
for i=round((10-1)*rand(1,1)+1);
(i,j)=4;
end;
end; cont=cont+1;
end;

%% Cálculo Diferencial;
a=round((5-1)*rand(10,200)+1);
for cont=0:1400;
for j=round((200-1)*rand(1,1)+1);
for i=round((10-1)*rand(1,1)+1);
(i,j)=4;
end;
end; cont=cont+1;
end;

%% Métodos de Ensino de Física 1;
a=round((5-1)*rand(10,200)+1);
for cont=0:1000;
for j=round((200-1)*rand(1,1)+1);
```

```

for i=round((10-1)*rand(1,1)+1);
(i,j)=4;
end;
end; cont=cont+1;
end;

%% Estudo Sócio-Político da Educação;
a=round((5-1)*rand(10,200)+1);
for cont=0:1400;
for j=round((200-1)*rand(1,1)+1);
for i=round((10-1)*rand(1,1)+1);
(i,j)=3;
end;
end; cont=cont+1;
end;

% Montando uma tensor tridimensional;
B=[{a,b,c,d}];
for k=(1:4);
    Y(k, :, :) = B{k};;
end;

```

## A.1 Código Utilizado para obter as Componentes do PCA

---

```

%Comando do Matlab para obter as componentes;
[coefs,score] = princomp(x);
% Plotando os gráficos do PCA;
% Plotrando a Matriz de Coeficiente;
scatter(coefs(:,1),coefs(:,2),'*');
for i = 1:10;
    text(coefs(i,1)+0.1,coefs(i,2)+0.01, abDoCoefs{i},'HorizontalAlignment','center');
end;

[coefs,score] = princomp(x);
% Plotando os gráficos do PCA;
% Plotrando a Matriz de Coeficiente;
plot(score(:,1),score(:,2),'s','MarkerEdgeColor','k','MarkerFaceColor','k','MarkerSize',10);

```



```
for i = 1:4;
    text(score(i,1)+0.1,score(i,2)+0.01, ab{i},'HorizontalAlignment','center');
end;
```

# Anexo **B**

## Código MATLAB

### B.0.1 Código Utilizado na Decomposição PARAFAC

```
clc;
clear all;
close all;

% Carregando a base de dados e ordenando como um tensor.
X = Y;

% Estabelecendo o número de componentes das matrizes A, B e C.
Fac = 2;

% Colocando opções no algoritmo (restrição, num de inter...).
Options(1) = 1e-6;
Options(2) = 0;
Options(3) = 1;
Options(4) = 0;
Options(5) = 10;
Options(6) = 2500;

% Colocando restrições nas matrizes A, B e C.
const = [2 1 1];
```

---

```
% Parâmetro utilizado caso você já conheça algumas das matrizes estimadas.
OldLoad = [];

% Estimaremos todos os modos das minhas matrizes
FixMode = [1 1 1];

% Nenhuma matriz de pesos será considerada.
Weights = [];

% Aplicando a decomposição PARAFAC COMPLETA.
[Factors,it,err,corcondia] = parafac(X,Fac,Options,const,OldLoad,FixMode,Weights)

% Analisando a consistência do núcleo.
Consistency = corcond(X,Factors,Weights,1);

% Criando as matrizes, A, B e C.
[A, B, C] = fac2let(Factors)
```

# Referências Bibliográficas

ALMEIDA, A. L. F. de. *Tensor Modeling and Signal Processing*. Tese (Doutorado) — Universidade Federal do Ceará, 2007.

BOLDRIN, J. L. *Algebra Linear*. 3. ed. Harba: CENGAGE Learning, 1986. ISBN 978-85-221-0744-5.

CICHOCKI, A. et al. *Nonnegative Matrix and Tensor Factorizations - Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*. shghgh: John Wiley & Sons Ltd, 2009. 477 p.

KOLDA, T. G.; BADER, B. W. Tensor decomposition and applications. *Society for Industrial and Applied Mathematics*, p. 46, 2009.

MARSH, H. W. The use of path analysis to estimate teacher and course effects in student ratings of instructional effectiveness. *Applied Psychological Measurement*, n. 6, p. 47–60, 1981.

MARSH, H. W.; DUNKIN., M. J. Students' evaluations of university teaching: A multidimensional perspective., journal = In Higher Education: Handbook of Theory and Research, year= 1992, keywords = Edited by J. Smart,.

MARSH, H. W.; DUNKIN, M. J. The multidimensionality os students' evaluations of teaching effectiveness: The generality of factor structures across academic discipline, instructor level, and course level. *Teaching and Teacher Education*, v. 1, p. 9–18, 1991.

MINGOTI, S. A. *Análise de dados através de métodos de estatísticas multivariada: uma abordagem aplicada*. 2. ed. Belo Horizonte: Ed. UFMG, 2005. ISBN 85-7041-451-X.

MURRAY H. G., J. P. R. Teacher personality traits and student instructional ratings in six types of university courses. *Journal of Educational Psychology*, n. 82, p. 250–261, 1990.

PEREYRA; SCHERER, G. Efficient computer manipulation of tensor products with application to multidimensional aproximation. *Math Comp*, 1973.

SCRIVEN, M. *Evaluation Ideologies. Evaluation Models-viewpoints on educational and human services evaluation*. [S.l.]: Boston, Kluwer-Nijhoff Publishing, 1993.

SMILDE, A.; BRO, R.; GELADI, P. *Multi-way Analysis - Applications in the Chemical Sciences*. [S.l.]: John Wiley & Sons Ltd, 2004. 381 p.

STEVENS, J. Applied multivariate statistics for the social sciences. *Harper and Now*, v. 1, n. 2, 1986.

STRANG, G. *álgebra linear e suas aplicações*. 4. ed. Massachusetts Institute of Technology: CENGAGE Learning, 2012. Tradução da 4ª Edição norte-americana. ISBN 978-85-221-0744-5.

STUDENTS' Evaluations of University Teaching: Dimensionality, Reliability, Validity, Potential Biases, and Utility. *Journal of Educational Psychology*, n. 76, p. 707–754, 1984.

STUDENTS' Evaluations of University Teaching: Research Findings, Methodological Issues, and Directions for Future Research. *Internacional Journal of Educational Research*, n. 11, p. 253–388, 1987.

STUFFLEBEAM. *The CIPP Models-viewpoints on Educational Evaluation*. [S.l.]: Boston, Kluwer-Nijhoff Publishing, 1993.