

EDGARD PEDROSO

**UMA CONTRIBUIÇÃO AO ESTUDO DO
CÁLCULO DO AUTOSISTEMA DA MATRIZ QUE
APARECE NO PROBLEMA DE ANÁLISE DE
CORRELAÇÃO CANÔNICA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção e Sistemas da Pontifícia Universidade Católica do Paraná, como requisito parcial para obtenção do título de Mestre em Engenharia de Produção e Sistemas.

Área de concentração: Gerência de produção e Logística
Orientador: Prof. Dr. Raimundo José Borges de Sampaio

Curitiba
[2002]

EDGARD PEDROSO

**UMA CONTRIBUIÇÃO AO ESTUDO DO
CÁLCULO DO AUTOSISTEMA DA MATRIZ QUE
APARECE NO PROBLEMA DE ANÁLISE DE
CORRELAÇÃO CANÔNICA**

Curitiba
[2002]

Pedroso, Edgard

Uma Contribuição ao Estudo do Cálculo do Autosistema da Matriz que Define o Problema da Análise de Correlação Canônica.

Curitiba, 2002.

Dissertação (Mestrado) - Pontifícia Universidade Católica do Paraná.
Programa de Pós-Graduação em Engenharia de Produção e Sistemas.

1. Análise de Correlação Canônica 2. Cálculo de autovalor e autovetor

Dedico este trabalho aos meus pais Edgard e Norma
e às minhas filhas Raphaela e Caroline

Agradecimentos

Agradeço ao Prof. Dr. Raimundo José Borges de Sampaio pelo apoio e orientação, sem os quais este trabalho não teria sido realizado.

Agradeço a todos os Professores do Curso de Mestrado pela dedicação nas aulas ministradas.

Agradeço ao Prof. Dr. Rubens Robles Ortega Júnior pela motivação dada para que eu fizesse o Curso de Mestrado.

Agradeço aos Professores Dr. Anselmo Chaves Neto, Jorge Assade Leludak, Anderson Wolupeck e Rosa Bertila Pizzato, e aos Senhores Lauro Deconto e Dornelles Vissotto Júnior, pela ajuda prestada.

Agradeço às minhas filhas, Raphaela e Caroline, pela compreensão e paciência que tiveram, quando de minha ausência em virtude do trabalho.

Agradeço a minha namorada Margarete pela ajuda, carinho e compreensão.

Agradeço a todos amigos e familiares que de alguma forma colaboraram para a realização deste trabalho.

Enfim, agradeço a Deus por tudo que me tem proporcionado.

Sumário

Agradecimentos	viii
Sumário	ix
Resumo	xii
Abstract	xiii
1 Introdução	1
1.1 Descrição do Problema	1
1.2 Objetivo	1
1.3 Justificativa	2
1.4 Metodologia	3
1.5 Organização	3
2 Revisão de Literatura	5
2.1 Tópicos da Álgebra Linear	5
2.1.1 Normas de Vetores	5
2.1.2 Erros Absoluto e Relativo	6
2.1.3 Normas de Matrizes e de Transformações Lineares	7
2.1.4 Inversas de Matrizes Perturbadas	10
2.1.5 Número de Condição de Matriz	11
2.1.6 Condição de Autosistemas	14
2.1.7 Decomposição Espectral	20

2.1.8	Decomposição de Valor Singular	21
2.1.9	Decomposição de Valor Singular Reduzida	24
2.1.10	Inversa Generalizada ou Pseudo-Inversa	25
2.1.11	Formas Quadráticas e Autovalores.....	26
2.1.12	Raiz Quadrada de Uma Matriz	28
2.2	Análise de Correlação Canônica.....	30
2.2.1	Vetores e Matrizes Aleatórias	30
2.2.2	Vetor de Médias, Matriz de Covariâncias e Matriz de Correlação.	31
2.3	Correlação Canônica	34
2.3.1	Introdução	34
2.3.2	Variáveis Canônicas e Correlações Canônicas	34
3	Material e Método	51
3.1	Um Novo Método Numérico para o Cálculo de Autovalor e Autovetor de Produto de Matrizes	51
3.1.1	Introdução	51
3.1.2	Caracterização da Matriz Produto (m)	53
3.1.3	Construção da Matriz de Cinco Blocos (M).....	53
3.1.4	Invertibilidade de M	57
3.1.5	Autosistema Relacionado com a Matriz M	58
3.1.6	Decomposição de Valor Singular de M	59
3.1.7	Redução do Produto $m = m_1.m_2.m_3.m_4.m_5$ ao Produto $m = Bm_3B^T$. 61	
3.1.8	Construção da Matriz de Três Blocos (W).....	62
3.1.9	Redução do Produto $m = m_1.m_2.m_3.m_4.m_5$ ao Produto $m = AA^T$	63
3.1.10	Construção da Matriz de Dois Blocos (N)	63

3.1.11	Decomposição de Valor Singular da Matriz A .	64
3.1.12	Relação entre o Autosistema de N e a Decomposição de Valor Singular de A .	65
3.2	Experimentos Numéricos	67
3.2.1	Geração das Matrizes m_i .	67
3.2.2	Cálculo do Autovalor e Autovetor da Matriz $m = m_1 m_2 m_3 m_4 m_5$.	67
3.2.3	Construção das Matrizes de Blocos, M , W e N .	68
3.2.4	Cálculo dos Autovalores e Autovetores das Matrizes M , W e N .	68
3.2.5	Cálculo do Autovalor e Autovetor da Matriz $m = AA^T$ pela decomposição singular de A .	68
3.2.6	Perturbação das Matrizes	69
3.2.7	Precisão Numérica.	69
3.2.8	Análise dos Resultados Experimentais	70
4	Conclusões	72
4.1	Propostas para Trabalhos Futuros	74
	Referências Bibliográficas	75
	A Rotina Para os Experimentos Numéricos	77
	B Resultados Experimentais Matrizes 5x5	91
	C Resultados Experimentais Matrizes 10x10	99
	D Resultados Experimentais Matrizes 15x15	105
	E Resultados Experimentais Matrizes 20x20	108

Resumo

Em várias situações precisa-se calcular os autovalores e os autovetores de uma matriz dada por um produto de matrizes. Um exemplo particular desse problema ocorre numa situação muito especial no âmbito da *Análise Multivariada*, quando precisa-se medir a relação linear entre dois conjuntos de variáveis aleatórias. Esse problema é denominado, no campo da Estatística, de *Análise de Correlação Canônica*, e é formulado como um problema de determinação do autosistema de um produto de cinco matrizes, $m = m_1 m_2 m_3 m_4 m_5$, onde m é uma matriz simétrica e definida positiva. Como este trabalho foi motivado pelo problema de análise de correlação canônica, será considerado aqui o problema de achar o autosistema da matriz $m = m_1 m_2 m_3 m_4 m_5$, onde m é simétrica e definida positiva. Assumi-se também que todas as matrizes são quadradas não singulares, $m_5^T = m_1$, $m_4^T = m_2$ e m_3 é simétrica e definida positiva. Como é sabido da literatura, se m é uma matriz simétrica e positiva definida, então o problema de calcular os autovalores é bem-condicionado. Infelizmente não se pode afirmar o mesmo quando se trata do problema de calcular os autovetores. De fato pode-se mostrar que até mesmo para matrizes diagonais, a introdução de pequenas perturbações podem resultar uma grande variação em seus autovetores. Isto ocorre frequentemente com autovetores associados com autovalores com algum tipo de agrupamento. Uma idéia, objetivo dessa dissertação, que parece ser útil para contornar esse problema de mal-condicionamento é calcular o autovalor e o autovetor associado, sem o resultado da multiplicação indicada em $m = m_1 m_2 m_3 m_4 m_5$. Experimentos numéricos aleatórios serão apresentados ao final.

Abstract

In many situations we need to calculate the eigenvalues and eigenvectors of a matrix which are given as a product of matrices. For instance, this happens, in particular, in a very especial situation in the domain of multivariate analysis, when we need to measure the linear relationship between two sets of random variables. This problem, that is called in the field of statistic as the canonical correlation problem, is formulated as the problem of finding the eigensystem of a product of five matrices, $m = m_1 m_2 m_3 m_4 m_5$ where m is symmetric and positive definite.

Since the analysis that follows has been motivated by the problem of canonical correlation, we are going to consider here the problem of finding the eigensystem of a matrix $m = m_1 m_2 m_3 m_4 m_5$ where m is symmetric and positive definite. We also assume that all the matrices are square nonsingular, $m_5^T = m_1$, $m_4 = m_2^T$, and m_3 is symmetric and positive definite.

It is well known from the literature that if m is symmetric and positive definite, then the condition number of the problem of finding an eigenvalue can be bounded, that is, the problem is well-conditioned. Unfortunately, we do not have the same result for the corresponding eigenvector problem. In fact, it can be shown that even for diagonal matrix, a little perturbation in the entries of matrix can result in a large variation of its eigenvectors. This happens frequently with eigenvectors related to eigenvalues with some kind of clustering. An idea that seems to be useful to get around this problem of ill-conditioning is to

calculate the eigenvalue and the eigenvector associated without performing the indicated multiplication in $m = m_1 m_2 m_3 m_4 m_5$.

In this work we define a 5-block matrix M , a 3-block matrix W and a 2-block matrix N , whose eigensystems are related to the eigensystem of m and then we calculate the eigensystem of m using the eigensystems of M , of W or of N . Furthermore, we reduce $m = m_1 m_2 m_3 m_4 m_5$ to $m = AA^T$, where $A = m_1 m_2 m_3^{1/2}$, and then, we find eigensystem of m using the singular value decomposition of A . Numerical random experiments will be reported in the end.

Capítulo 1 - Introdução

1.1 Descrição do Problema

Em diversas situações reais tem-se o problema de determinar o autosistema de uma matriz dada por um produto de matrizes. Um exemplo particular desse problema ocorre na *Análise Multivariada*, mais precisamente na *Análise de Correlação Canônica* que tem por objetivo identificar e quantificar as associações lineares entre dois conjuntos de variáveis aleatórias. A formulação dessa análise culmina com o problema de achar o autosistema de uma matriz, simétrica e definida positiva, resultante de um produto de cinco matrizes. Como este trabalho foi motivado pelo problema de análise de correlação canônica, será considerado aqui o problema de achar os autovalores e autovetores associados da matriz $m = m_1 m_2 m_3 m_4 m_5$, onde m é uma matriz simétrica e definida positiva. Será assumido também que todas as matrizes são quadradas não singulares, $m_5^T = m_1$, $m_4^T = m_2$ e que m_3 é simétrica e definida positiva.

1.2 Objetivo

O objetivo desse trabalho é expor e contornar o problema de mal-condicionamento do cálculo dos autovetores, determinando o autosistema de m , sem o resultado da multiplicação indicada $m = m_1 m_2 m_3 m_4 m_5$. Será definida uma matriz M de cinco blocos, uma matriz W de três blocos e uma matriz N de dois blocos, cujos autosistemas estão

relacionados com o autosistema m e, então, calcula-se o autosistema de m , usando os autosistemas de M, W e N . Além disso, o cálculo de m será reduzido para $m = AA^T$, com $A = m_1 m_2 m_3^{1/2}$, e o autosistema de m será calculado através da decomposição de valor singular da matriz A .

1.3 Justificativa

Na Área de Gerência de Produção e Logística, quando se aborda a otimização de Sistemas Produtivos, é freqüente depara-se com problemas que envolvem grande quantidade de dados. Assim, é provável que ocorram inexatidões na coleta desses dados, quer por pequenos erros de medida ou por arredondamentos. Uma vez introduzidos no método de cálculo, esses erros podem trazer resultados bem diferentes daqueles que seriam obtidos, caso os dados de entrada fossem de melhor qualidade, comprometendo a confiabilidade nos resultados. No caso específico deste trabalho, embora se saiba que o problema de achar os autovalores de matrizes simétricas e definidas positivas é sempre bem-condicionado, o mesmo não pode ser afirmado no que diz respeito ao problema de achar os autovetores, eis que pequenas perturbações inseridas na matriz podem fazer com que os autovetores calculados sejam bem diferentes dos reais. Além disso, o custo de processamento de grande quanti-

dade de dados é alto e, assim, deve-se sempre procurar processos de cálculo que demandem menor tempo de processamento.

1.4 Metodologia

Inicialmente foi feita uma pesquisa bibliográfica para se ter acesso aos principais e mais recentes resultados sobre o método de cálculo do autosistema da matriz que define o problema de correlação canônica. Após concluída a revisão bibliográfica, fez-se uma análise crítica dos métodos utilizados e, então, foram introduzidos novos métodos. Com a comprovação teórica dos novos métodos, elaborou-se uma estratégia para a experimentação numérica e então confrontaram-se os resultados.

1.5 Organização

O trabalho está organizada em quatro partes. No segundo capítulo, desenvolve-se uma Revisão de Literatura sobre *Tópicos da Álgebra Linear* necessários para a compreensão e desenvolvimento do objetivo principal do trabalho e aborda-se a *Análise da Correlação Canônica*, que serviu como motivação para o desenvolvimento do tema principal desta dissertação, pois aí surge um exemplo de aplicação do cálculo do autosistema da matriz que é dada pelo produto de cinco matrizes com características peculiares. Uma nova abordagem, para a resolução do problema de achar o autosistema daquela matriz, a qual é dada pelo produto, é então proposta no terceiro capítulo, na seção *Um Novo Método Numérico para o Cálculo de Autovalor e Autovetor de Produto de Matrizes*. Na seção *Experimentos Numéri-*

cos do terceiro capítulo é apresentada a elaboração dos experimentos numéricos aleatórios.

Ao final, são apresentados a conclusão e os resultados dos experimentos numéricos.

Capítulo 2 - Revisão de Literatura

2.1 Tópicos da Álgebra Linear

2.1.1 Normas de Vetores

A norma de um vetor em \mathbb{R}^n é uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$, representada como $f(x) = \|x\|$, onde:

$$(i) \quad \|x\| \geq 0 \quad \forall x \in \mathbb{R}^n \quad (\|x\| = 0 \Leftrightarrow x = 0),$$

$$(ii) \quad \|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in \mathbb{R}^n \quad (\text{desigualdade triangular}) \text{ e}$$

$$(iii) \quad \|\alpha x\| = |\alpha| \|x\| \quad \forall \alpha \in \mathbb{R}, \forall x \in \mathbb{R}^n.$$

Uma classe usual de norma de vetor são as normas- p definidas por

$$\|x\|_p = (|x_1|^p + \dots + |x_n|^p)^{1/p} \quad 1 \leq p < \infty.$$

Destas normas, as mais usadas são:

a norma-1:

$$\|x\|_1 = |x_1| + \dots + |x_n|;$$

a norma-2:

$$\|x\|_2 = (|x_1|^2 + \dots + |x_n|^2)^{1/2} = (x^T x)^{1/2}; \quad (1)$$

e a norma- ∞ :

$$\|x\|_{\infty} = \max_{1 \leq i \leq n} \{|x_i|\}.$$

Outras propriedades de interesse são:

- (iv) Desigualdade de Cauchy-Schwartz : $(x^T y) \leq \|x\|_2 \|y\|_2$.
- (v) Todas as normas no \mathbb{R}^n são equivalentes, isto é, se $\|\cdot\|_{\alpha}$ e $\|\cdot\|_{\beta}$ são normas em \mathbb{R}^n , então existem constantes positivas, c_1 e c_2 tal que $c_1 \|x\|_{\alpha} \leq \|x\|_{\beta} \leq c_2 \|x\|_{\alpha}$ para todo $x \in \mathbb{R}^n$.

2.1.2 Erros Absoluto e Relativo

Supondo que $\hat{x} \in \mathbb{R}^n$ seja uma aproximação de $x \in \mathbb{R}^n$, então o erro absoluto cometido ao se representar x por \hat{x} é dado por:

$$\epsilon_{abs} = \|\hat{x} - x\|$$

e, se $x \neq 0$, o erro relativo correspondente será:

$$\epsilon_{rel} = \frac{\|\hat{x} - x\|}{\|x\|}.$$

O erro relativo na norma- ∞ pode ser traduzido em uma relação sobre o número de dígitos significativos corretos em \hat{x} . Em particular, se

$$\frac{\|\hat{x} - x\|_{\infty}}{\|x\|_{\infty}} \approx 10^{-p}$$

então a maior componente de \hat{x} tem aproximadamente p dígitos significativos corretos.

2.1.3 Normas de Matrizes e de Transformações Lineares

Seja \mathcal{A} uma transformação linear de um espaço vetorial V com uma norma $\|\cdot\|_V$ em um espaço vetorial W com uma norma $\|\cdot\|_W$. Para cada vetor não-nulo v em V , o quociente

$$\frac{\|\mathcal{A}(v)\|_W}{\|v\|_V}$$

mede a ampliação causada pela transformação \mathcal{A} . Verifica-se que, em um espaço qualquer V de dimensão *finita*, este quociente não se torna arbitrariamente grande e atinge seu valor máximo para um determinado vetor v . Sendo assim, o referido quociente pode ser limitado pelo número correspondente àquele valor máximo atingido, o qual denominamos de *supremo*. Se o *supremo* for de valor baixo, então todo vetor v terá seu tamanho reduzido por \mathcal{A} , enquanto que se o *supremo* for de valor alto, alguns vetores v terão seus tamanhos muito aumentados por \mathcal{A} . Por analogia com as normas de vetores que medem o tamanho de um vetor, denomina-se este número, o *supremo*, de *norma da transformação \mathcal{A}* , ou seja,

$$\|\mathcal{A}\|_{V,W} = \sup_{v \neq 0} \left\{ \frac{\|\mathcal{A}(v)\|_W}{\|v\|_V} \right\}.$$

PROPRIEDADES:

Sejam V, W e Z espaços vetoriais e \mathcal{A} , \mathcal{B} e \mathcal{C} transformações lineares de V em W , de V em W e de W em Z , respectivamente, então:

- (i) $\|\mathcal{A}(v)\|_W \leq \|\mathcal{A}\|_{V,W} \|v\|_V$,
- (ii) $\|\mathcal{A}\|_{V,W} \geq 0$ sendo $\|\mathcal{A}\|_{V,W} = 0$ se e somente se $\mathcal{A}(v) = 0$ para todo v ,
- (iii) $\|k \mathcal{A}\|_{V,W} = |k| \|\mathcal{A}\|_{V,W}$ para todos os escalares k ,
- (iv) $\|\mathcal{A} + \mathcal{B}\|_{V,W} \leq \|\mathcal{A}\|_{V,W} + \|\mathcal{B}\|_{V,W}$,
- (v) $\|\mathcal{C} \mathcal{B}\|_{V,Z} \leq \|\mathcal{C}\|_{W,Z} \|\mathcal{A}\|_{V,W}$ e
- (vi) $\|\mathcal{I}\|_{V,V} = 1$, onde $\mathcal{I}(v) = v$ para todo v em V .

Aqui trata-se do espaço vetorial \mathbb{R}^n e das transformações lineares \mathcal{A} definidas pelo produto matricial $\mathcal{A}(x) = Ax$. Assim, são consideradas as normas de transformações induzidas pelas normas de vetores $\|\cdot\|_1$, $\|\cdot\|_2$ e $\|\cdot\|_\infty$ sobre as transformações lineares definidas por matrizes.

Seja a matriz $A_{m \times n}$ e a transformação linear de \mathbb{R}^n em \mathbb{R}^m definida por $\mathcal{A}(x) = Ax$. Representando-se por $\|A\|_1$, $\|A\|_2$ e $\|A\|_\infty$ as normas correspondentes da transformação \mathcal{A} induzidas e usando as normas vetoriais apropriadas tanto para o domínio \mathbb{R}^n como para o contradomínio \mathbb{R}^m , é possível demonstrar que :

$$\|A\|_1 = \max_{x \neq 0} \left\{ \frac{\|Ax\|_1}{\|x\|_1} \right\} = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|, \quad (2)$$

$$\|A\|_\infty = \max_{x \neq 0} \left\{ \frac{\|Ax\|_\infty}{\|x\|_\infty} \right\} = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}| \quad \text{e} \quad (3)$$

$$\|A\|_2 = \max_{x \neq 0} \left\{ \frac{\|Ax\|_2}{\|x\|_2} \right\} \quad \text{ou}$$

$$\|A\|_2 = [\text{máx. autovalor de } (A^T A)]^{1/2} = \text{valor singular máx. de } A. \quad (4)$$

Como este último resultado é mais usual, será feita a seguir sua comprovação. Seja x o autovetor de $A^T A$ associado a seu maior autovalor λ_{\max} . Então,

$$\|A\|_2^2 = \max_{x \neq 0} \left\{ \frac{\|Ax\|_2^2}{\|x\|_2^2} \right\} = \max_{x \neq 0} \left\{ \frac{x^T A^T A x}{x^T x} \right\} = \frac{x^T (\lambda_{\max}) x}{x^T x} = \lambda_{\max},$$

e nenhum outro x produzirá um valor maior para esta razão, como será mostrado. A matriz simétrica $A^T A$ possui um conjunto $\{q_1, q_2, \dots, q_n\}$ de autovetores *ortonormais* e todo vetor x pode ser escrito como combinação linear desses autovetores. Assim,

$$\frac{x^T A^T A x}{x^T x} = \frac{(c_1 q_1 + \dots + c_n q_n)^T (c_1 \lambda_1 q_1 + \dots + c_n \lambda_n q_n)}{(c_1 q_1 + \dots + c_n q_n)^T (c_1 q_1 + \dots + c_n q_n)}$$

e como $q_i^T q_j = 0$, então,

$$\frac{x^T A^T A x}{x^T x} = \frac{c_1^2 \lambda_1 + \dots + c_n^2 \lambda_n}{c_1^2 + \dots + c_n^2}.$$

Esta razão é conhecida como *Quociente de Rayleigh*. O Quociente de Rayleigh atinge seu valor máximo, igual a λ_{\max} , quando todos os coeficientes c são iguais a zero exceto aquele que multiplica λ_{\max} e atinge seu valor mínimo, igual a λ_{\min} , quando todos os coeficientes c são iguais a zero exceto aquele que multiplica λ_{\min} . Portanto, para qualquer vetor $x \neq 0$, obtém-se

$$\lambda_{\min} \text{ de } (A^T A) \leq \frac{x^T A^T A x}{x^T x} \leq \lambda_{\max} \text{ de } (A^T A). \quad (5)$$

Outra norma de matriz de uso comum na álgebra linear é a norma de Frobenius. A norma de Frobenius *não é induzida por norma de vetor* e sua definição é dada pela expressão

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} \quad (6)$$

onde a_{ij} representa o elemento da i – *ésima* linha e j – *ésima* coluna da matriz A .

2.1.4 Inversas de Matrizes Perturbadas

Sabe-se o quão importante é perceber que, em virtude da inexatidão de dados, pode-se trabalhar com vetores ou matrizes que sofreram pequenas perturbações de seus valores exatos. Um dos problemas que a introdução desses erros pode causar é que uma matriz perturbada $(A + \delta A)$ seja singular, quando de fato a matriz verdadeira A não o é. Para pequenas perturbações, pode-se identificar estas possibilidades, conforme será visto no lema a seguir.

LEMA DE BANACH : Seja P uma matriz $n \times n$ e $\|\cdot\|$ uma das normas de operadores.

Se $\|P\| < 1$, então $I + P$ é não-singular e

$$\frac{1}{1 + \|P\|} \leq \|(I + P)^{-1}\| \leq \frac{1}{1 - \|P\|}. \quad (7)$$

DEMONSTRAÇÃO [NOB96]: Sabe-se que $I + P$ é não singular se e somente se a única solução para $(I + P)x = 0$ é $x = 0$. Agora supondo que $(I + P)x = 0$ com $x = -Px$, neste caso, tem-se $\|x\| = \|-Px\| = \|Px\| \leq \|P\| \|x\|$, mas como $\|P\| < 1$, isto é uma contradição, a menos que x seja nulo, e portanto, $(I + P)^{-1}$ existe. Ainda, se $(I + P)^{-1} = B$, tem-se $I = B(I + P) = B + BP$ e

$$1 = \|I\| = \|B(I + P)\| \leq \|B\| \|I + P\| \leq \|B\| (1 + \|P\|)$$

então

$$\frac{1}{1 + \|P\|} \leq \|B\|.$$

Como $B = I - BP$ segue também que

$$\|B\| = \|I - BP\| \leq 1 + \|BP\| \leq 1 + \|B\| \|P\|$$

e assim

$$\|B\| \leq \frac{1}{1 - \|P\|}.$$

TEOREMA : Sejam A e δA matrizes $n \times n$ com A não singular e $\|\cdot\|$ uma das normas de operadores. Se $\|(\delta A) A^{-1}\| < 1$ ou $(\|A^{-1}(\delta A)\| < 1)$, de maneira particular $(\|\delta A\| < \|A^{-1}\|^{-1})$, então $A + \delta A$ é não singular e

$$\|(A + \delta A)^{-1}\| \leq (1 - \|(\delta A) A^{-1}\|)^{-1} \|A^{-1}\|. \quad (8)$$

DEMONSTRAÇÃO [NOB96] : Como A é não singular pode-se escrever $A + \delta A = (I + P) A$, com $P = (\delta A) A^{-1}$. Pelo *lema de Banach* sabemos que $(I + P)$ é não singular e

$$\|(I + P)^{-1}\| \leq (1 - \|P\|)^{-1} \text{ ou } \|(I + P)^{-1}\| \leq (1 - \|(\delta A) A^{-1}\|)^{-1},$$

como

$$[A(I + P)]^{-1} = (I + P)^{-1} A^{-1},$$

então

$$\|[A(I + P)]^{-1}\| \leq \|(I + P)^{-1}\| \|A^{-1}\|$$

e

$$\|[A(I + P)]^{-1}\| = \|(A + \delta A)^{-1}\| \leq (1 - \|(\delta A) A^{-1}\|)^{-1} \|A^{-1}\|.$$

2.1.5 Número de Condição de Matriz

Como é de conhecimento, erros de medida ou arredondamento podem trazer sérios problemas na solução de sistemas lineares. Alguns sistemas são mais sensíveis à introdução desses erros (*perturbações*) que outros. Uma maneira de medir essa sensibilidade

é determinar o *número de condição* da matriz do sistema, sobre o qual versa o teorema a seguir.

TEOREMA : Seja A uma matriz $m \times n$ e seja $\|\cdot\|$ uma das normas de operadores e sua norma vetorial correspondente. Sejam A não-singular e x uma solução de $Ax = b$. Supondo uma perturbação da matriz A , mudando-a para $A + \delta A$ e uma perturbação no vetor b mudando-o para $b + \delta b$ e que essas perturbações façam x mudar para $x + \delta x$. Considerando ainda, $\|(\delta A) A^{-1}\| < 1$. Então,

$$\frac{\|\delta x\|}{\|x\|} \leq M \cdot c(A) \cdot \left(\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta A\|}{\|A\|} \right) \quad (9)$$

onde

$$M = [1 - \|(\delta A) A^{-1}\|]^{-1} \quad (10)$$

e

$$c(A) = \|A\| \|A^{-1}\| \text{ é o número de condição.} \quad (11)$$

DEMONSTRAÇÃO[NOB96]: Considerando que $\|(\delta A) A^{-1}\| < 1$ e o teorema enunciado sobre as *inversas de matrizes perturbadas*, podemos afirmar que $A + \delta A$ é não-singular, assegurando dessa forma que a solução $x + \delta x$ do problema perturbado existe.

Agora, introduzindo erros na matriz A e no vetor b , tem-se

$$(A + \delta A)(x + \delta x) = b + \delta b,$$

que produz

$$(A + \delta A) \delta x = b + \delta b - Ax - \delta Ax = \delta b - \delta Ax,$$

então

$$\delta x = (A + \delta A)^{-1} (\delta b - \delta Ax)$$

e

$$\|\delta x\| \leq \|(A + \delta A)^{-1}\| \|(\delta b - \delta Ax)\|.$$

Mas pelo teorema enunciado sobre as *inversas de matrizes perturbadas*

$$\|(A + \delta A)^{-1}\| \leq [1 - \|(\delta A) A^{-1}\|]^{-1} \|A^{-1}\|$$

ou no presente caso

$$\|(A + \delta A)^{-1}\| \leq M \|A^{-1}\|.$$

Portanto, pode-se escrever

$$\|\delta x\| \leq M \cdot \|A^{-1}\| (\|\delta b\| + \|\delta A\| \|x\|)$$

que dividindo por $\|A\| \|x\|$ resulta $\frac{\|\delta x\|}{\|x\|} \leq M \cdot \|A\| \|A^{-1}\| \cdot \left(\frac{\|\delta b\|}{\|A\| \|x\|} + \frac{\|\delta A\|}{\|A\|} \right)$. Como

$Ax = b$ e $\|A\| \|x\| \geq \|b\|$ então finalmente obtém-se

$$\frac{\|\delta x\|}{\|x\|} \leq M \cdot c(A) \cdot \left(\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta A\|}{\|A\|} \right).$$

Se a perturbação δA de A for suficientemente pequena, a constante M estará próxima de 1 e portanto a mudança global $\frac{\|\delta x\|}{\|x\|}$ nas incógnitas, devida às pequenas perturbações introduzidas em A e em b , será pequena, se $c(A)$ não for demasiadamente grande. Portanto, para um valor de $c(A)$ moderado, as equações serão *bem-condicionadas*. Porém um valor de $c(A)$ grande não implica necessariamente que as equações sejam *mal-condicionadas*.

Outra forma de observar o significado do número de condição é através do teorema enunciado sobre as *inversas de matrizes perturbadas*. Como foi visto, se A é não-singular então $A + \delta A$ é não-singular para todo δA com $\|\delta A\| < \|A^{-1}\|^{-1}$, ou seja, com

$$\frac{\|\delta A\|}{\|A\|} < \frac{1}{\|A\| \|A^{-1}\|}.$$

Portanto $A + \delta A$ é *não-singular* sempre que

$$\frac{\|\delta A\|}{\|A\|} < \frac{1}{c(A)}$$

ou de outra forma, se $A + \delta A$ é *singular*, então

$$\frac{\|\delta A\|}{\|A\|} \geq \frac{1}{c(A)}$$

resultando uma cota inferior para $c(A)$. Na verdade pode ser mostrado que

$$\frac{1}{c(A)} = \min \left\{ \frac{\|\delta A\|}{\|A\|} \text{ tal que } A + \delta A \text{ é singular} \right\}. \quad (12)$$

Logo $c(A)$ é grande, se e somente se existe uma matriz *singular* $A + \delta A$ com $\frac{\|\delta A\|}{\|A\|}$ pequeno.

2.1.6 Condição de Autosistemas

Aqui descreve-se o problema do tamanho das perturbações dos autovalores e autovetores causado por perturbações em A como sendo um problema de *condição* dos autovalores e autovetores de A . Assim será visto o importante teorema a seguir.

TEOREMA DO CÍRCULO DE GERSHGORIN: Cada autovalor λ de uma matriz $B_{n \times n}$ satisfaz pelo menos uma das desigualdades

$$|\lambda - b_{ii}| \leq r_i, \text{ onde } r_i = \sum_{\substack{j=1 \\ j \neq i}}^n |b_{ij}| \quad (i = 1, \dots, n). \quad (13)$$

Ou seja, cada autovalor pertence a pelo menos um dos discos com centro b_{ii} e raio r_i no plano complexo (figura 1, Círculo de Gershgorin).

DEMONSTRAÇÃO [NOB96]: Seja λ um autovalor da matriz $B_{n \times n}$ e $x_{n \times 1}$ o autovetor associado, então $Bx = \lambda x$ e $(B - \lambda I)x = 0$, ou seja,

$$\begin{pmatrix} b_{11} - \lambda & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} - \lambda & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

que produz

$$\begin{cases} (b_{11} - \lambda)x_1 + b_{12}x_2 + \cdots + b_{1n}x_n = 0 \\ b_{21}x_1 + (b_{22} - \lambda)x_2 + \cdots + b_{2n}x_n = 0 \\ \vdots + \vdots + \cdots + \vdots = \vdots \\ b_{n1}x_1 + b_{n2}x_2 + \cdots + (b_{nn} - \lambda)x_n = 0 \end{cases}$$

ou

$$\begin{cases} b_{12}x_2 + \cdots + b_{1n}x_n = (\lambda - b_{11})x_1 \\ b_{21}x_1 + b_{2n-1}x_{n-1} + b_{2n}x_n = (\lambda - b_{22})x_2 \\ \vdots + \vdots + \cdots + \vdots = \vdots \\ b_{n1}x_1 + \cdots + b_{n-1}x_{n-1} = (\lambda - b_{nn})x_n \end{cases}.$$

Somando as equações, obtém-se

$$\sum_{\substack{j=1 \\ j \neq i}}^n b_{ij} x_j = (\lambda - b_{ii}) x_i \quad (i = 1, \dots, n).$$

Agora supondo que x_k é o maior elemento do autovetor x , então $\left| \frac{x_j}{x_k} \right| \leq 1$ para todo j e pode-se escrever

$$(\lambda - b_{kk}) x_k = \sum_{\substack{j=1 \\ j \neq k}}^n b_{kj} x_j$$

que leva a

$$|\lambda - b_{kk}| |x_k| \leq \sum_{\substack{j=1 \\ j \neq k}}^n |b_{kj}| |x_j|,$$

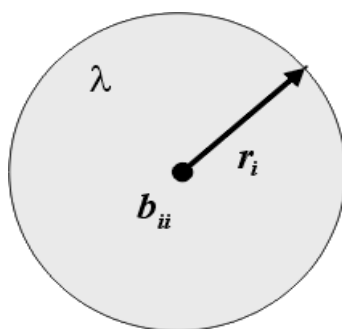
dividindo ambos os membros por $|x_k|$, obtém-se

$$|\lambda - b_{kk}| \leq \sum_{\substack{j=1 \\ j \neq k}}^n |b_{kj}| \left| \frac{x_j}{x_k} \right| \leq \sum_{\substack{j=1 \\ j \neq k}}^n |b_{kj}|$$

ou

$$|\lambda - b_{ii}| \leq r_i \text{ com } r_i = \sum_{\substack{j=1 \\ j \neq i}}^n |b_{ij}|, \quad (i = 1, \dots, n),$$

como ilustrado a seguir.



1. Círculo de Gershgorin

Além disso, é possível provar que se a união de p discos for disjunta dos restantes, então há precisamente p autovalores de B nesta união dos p discos.

Para um estudo preliminar da condição de auto-sistemas, analisaremos um caso muito particular, porém ilustrativo, quando A é uma matriz diagonal. Seja a matriz

$$A = \begin{pmatrix} a_{11} & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & a_{nn} \end{pmatrix},$$

então seus autovalores serão $\lambda_i = a_{ii}$ para $i = 1, 2, \dots, n$, com os autovetores correspondentes iguais aos vetores coluna unitários e_i . Assim, o interesse é saber o que acontece quando se perturba a matriz A mudando-a para $A + \delta A$, sendo que δA representa uma matriz com *pequena* perturbação. Aplicando o teorema de *Gershgorin*, tem-se que os au-

tovalores da matriz perturbada $A + \delta A$ pertencerão aos discos

$$|(a_{ii} + \delta a_{ii}) - \lambda| \leq r_i = \sum_{\substack{j=1 \\ j \neq i}}^n |\delta a_{ij}|, \quad (i = 1, \dots, n),$$

mas como foi visto, os elementos diagonais a_{ii} de A são os próprios autovalores λ_i , ou seja, $a_{ii} = \lambda_i$, logo pode-se escrever

$$|(\lambda_i + \delta a_{ii}) - \lambda| \leq r_i \text{ e então } |\lambda_i - \lambda| \leq r_i + |\delta a_{ii}|, \quad (i = 1, \dots, n).$$

O valor de $r_i + |\delta a_{ii}|$ pode ser limitado pelo máximo das somas dos valores absolutos de linhas, que é a norma $\|\delta A\|_\infty$. E assim pode-se concluir que

$$|\lambda_i - \lambda| \leq \|\delta A\|_\infty, \quad (14)$$

e, portanto, *os autovalores de uma matriz diagonal são sempre bem-condicionados*. Já no que diz respeito aos autovetores de uma matriz diagonal, a análise é mais complexa. Os exemplos práticos nos mostram que podemos esperar que uma matriz diagonal com autovalores distintos e bem separados tenha autovetores *bem-condicionados*. Quando os autovalores são repetidos ou muito próximos entre si, espera-se que a matriz diagonal tenha autovetores *mal-condicionados*.

Agora, será analisado um caso mais complexo, o de uma matriz $n \times n$ que possua um conjunto de n autovetores independentes.

TEOREMA : Seja uma matriz $A_{n \times n}$ com o conjunto $\{x_1, \dots, x_n\}$ independente de autovetores e com os respectivos autovalores $\lambda_1, \dots, \lambda_n$. Supondo um número μ e um vetor unitário v ($\|v\| = 1$, onde $\|\cdot\|$ representa as normas $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$), como aproximações, respectivamente, de um autovalor e de um autovetor da matriz A e considerando ainda, r a

quantidade pela qual μ e v deixam de fazer parte do auto-sistema, ou seja,

$$Av = \mu v + r \quad \text{ou} \quad r = Av - \mu v. \quad (15)$$

Sendo $P = (x_1 \cdots x_n)$ a matriz dos autovetores, então pelo menos um dos autovalores λ_i de A satisfaz

$$|\lambda_i - \mu| \leq \|r\| \|P\| \|P^{-1}\|. \quad (16)$$

Se P é unitária (*em particular, se A é real simétrica ou hermitiana ou normal*), então pelo menos um dos autovalores λ_i de A satisfaz

$$|\lambda_i - \mu| \leq \|r\|_2. \quad (17)$$

DEMONSTRAÇÃO [NOB96]: Como A tem um conjunto linearmente independente de n autovetores, então $A = P\Lambda P^{-1}$, onde Λ é a matriz diagonal cujo i -ésimo elemento diagonal é λ_i . Portanto $r = Av - \mu v$ pode ser reescrito como

$$r = P\Lambda P^{-1}v - \mu v = (P\Lambda P^{-1} - I\mu)v = P(\Lambda P^{-1} - P^{-1}\mu)v$$

ou seja,

$$r = P(\Lambda - \mu I)P^{-1}v.$$

Se μ for um autovalor de A , é imediata a verificação do teorema, uma vez que para algum λ_i de A , $|\lambda_i - \mu| = 0$. Se μ não for um autovalor de A , então $(\Lambda - \mu I)$ é não singular e

$$v = P(\Lambda - \mu I)^{-1}P^{-1}r,$$

que produz

$$\|v\| \leq \|P\| \|(\Lambda - \mu I)^{-1}\| \|P^{-1}\| \|r\|$$

ou

$$1 \leq \|P\| \|(\Lambda - \mu I)^{-1}\| \|P^{-1}\| \|r\|.$$

Como $(\Lambda - \mu I)^{-1}$ é uma matriz diagonal, então

$$\|(\Lambda - \mu I)^{-1}\| = \max_i |\lambda_i - \mu|^{-1} = \left[\min_i |\lambda_i - \mu| \right]^{-1}$$

e pode-se escrever

$$\min_i |\lambda_i - \mu| \leq \|r\| \|P\| \|P^{-1}\|.$$

Logo $|\lambda_i - \mu| \leq \|r\| \|P\| \|P^{-1}\|$ se verifica para algum λ_i como pretendia-se demonstrar.

Além disso, se P é unitária, P^{-1} será unitária e $\|P\|_2 = \|P^{-1}\|_2 = 1$, e então $|\lambda_i - \mu| \leq \|r\|_2$, como também pretendia-se demonstrar. Este teorema pode ser utilizado para estudar a condição de autovalores, ou seja, se μ for tomado como um autovalor associado a um autovetor v da matriz perturbada $A + \delta A$, com $\|v\| = 1$, então

$$(A + \delta A)v = \mu v$$

$$Av - \mu v = -\delta Av$$

$$r = -\delta Av$$

$$\|r\| \leq \|\delta A\| \|v\|$$

$$\|r\| \leq \|\delta A\|.$$

Pelo teorema, para algum λ_i de A , vale $|\lambda_i - \mu| \leq \|r\| \|P\| \|P^{-1}\|$, então

$$|\lambda_i - \mu| \leq \|\delta A\| \|P\| \|P^{-1}\|.$$

Se P é unitária (em particular, se A é real simétrica ou hermitiana ou normal), então pelo menos um dos autovalores λ_i de A satisfaz

$$|\lambda_i - \mu| \leq \|\delta A\|.$$

E portanto os autovalores das matrizes hermitianas são sempre bem-condicionados.

2.1.7 Decomposição Espectral

Denomina-se espectro de uma matriz A ao conjunto de seus autovalores. A descrição do espectro feita a seguir é conhecida como *teorema espectral* para matrizes simétricas, que é demonstrado em [NOB96]. Sendo $A_{n \times n}$ uma matriz simétrica, então pode-se enunciar as seguintes propriedades:

- (i) A possui n autovalores *reais*, contando com as multiplicidades;
- (ii) O auto-espaço correspondente a cada autovalor λ tem dimensão igual à multiplicidade de λ como raiz da equação característica;
- (iii) Os autovetores correspondentes a autovalores distintos são ortogonais, bem como, os respectivos auto-espaços;
- (iv) A matriz A é diagonalizável por matriz ortogonal.

Agora, suponha $A = P \Lambda P^T$ tal que as colunas de $P_{n \times n} = [u_1 \cdots u_n]$ sejam formadas pelos autovetores de A , e que os autovalores correspondentes $\lambda_1, \dots, \lambda_n$, formem a diagonal principal da matriz diagonal Λ . Assim, pode-se escrever

$$A = P \Lambda P^T$$

$$A = (u_1 \cdots u_n) \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} u_1^T \\ \vdots \\ u_n^T \end{pmatrix}$$

$$A = (\lambda_1 u_1 \cdots \lambda_n u_n) \begin{pmatrix} u_1^T \\ \vdots \\ u_n^T \end{pmatrix}$$

ou

$$A = \lambda_1 (u_1 u_1^T) + \cdots + \lambda_n (u_n u_n^T), \quad (18)$$

que representa a *decomposição espectral* de A , isto porque A é decomposta em partes definidas por seu espectro (*autovalores*). Os termos $\lambda_i (u_i u_i^T)$, ($i = 1, \dots, n$), são matrizes $n \times n$ de posto 1, a exemplo da matriz $\lambda_1 (u_1 u_1^T)$, onde cada coluna é múltipla de u_1 . Observamos ainda que se $\|u_i\| = 1$, cada matriz $(u_i u_i^T)$ é uma matriz de projeção, eis que, para cada vetor x em \mathbb{R}^n tem-se $(u_i u_i^T) x$ como sua projeção sobre o subespaço gerado por u_i . A idéia que norteia a decomposição espectral pode ser estendida para matrizes retangulares. Neste caso, os vetores usados na expansão de A seriam os autovetores das matrizes AA^T e $A^T A$.

2.1.8 Decomposição de Valor Singular

A decomposição de valor singular é embasada na propriedade: “o valor absoluto de um autovalor λ de uma matriz simétrica B mede a quantidade que B estica ou encurta um determinado autovetor associado a λ ”. Se $Bx = \lambda x$ e $\|x\| = 1$, então

$$\|Bx\| = \|\lambda x\| = |\lambda| \|x\| = |\lambda|.$$

Sendo λ_1 o autovalor de maior módulo, então um autovetor unitário associado v_1 indicará a direção na qual o efeito de esticar de B é maior, ou seja, o comprimento de Bx é máximo quando $x = v_1$, e será de $\|B v_1\| = |\lambda_1|$. Uma descrição de v_1 e λ_1 , similar a essa, será feita a seguir para matrizes retangulares.

Seja uma matriz $A_{m \times n}$ (supondo $m > n$), então $A^T A$ é simétrica e pode ser diagonalizada por uma matriz ortogonal. Supondo $\{v_1, \dots, v_n\}$ uma base ortonormal para \mathbb{R}^n , consistindo nos autovetores de $A^T A$ e $\lambda_1, \dots, \lambda_n$, os autovalores associados. Então para $0 \leq i \leq n$, tem-se

$$\|A v_i\|^2 = (A v_i)^T (A v_i) = v_i^T A^T A v_i$$

como v_i é um autovetor de $A^T A$ pode-se escrever

$$\|A v_i\|^2 = v_i^T (\lambda_i v_i)$$

mas v_i é um vetor unitário, logo

$$\|A v_i\|^2 = \lambda_i$$

portanto os autovalores de $A^T A$ são todos *reais e não-negativos*. Assim, pode-se supor um ordenamento decrescente dos autovalores de $A^T A$, tal que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$, ou dos quadrados dos *valores singulares* de A , $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_n^2$, eis que, $\sigma_i = \sqrt{\lambda_i}$.

TEOREMA : Supondo $\{v_1, \dots, v_n\}$, uma base ortonormal para \mathbb{R}^n , consistindo em autovetores de $A^T A$, ordenados de tal forma que seus respectivos autovalores associados satisfaçam $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ e que A possua r valores singulares não-nulos.

Então $\{A v_1, \dots, A v_r\}$ é uma base ortonormal para o espaço coluna de A ($\mathcal{R}(A)$) e o posto de A é r .

DEMONSTRAÇÃO [LAY97]: Como v_i e $\lambda_j v_j$ são ortogonais para $i \neq j$, tem-se

$$(A v_i)^T (A v_j) = v_i^T A^T A v_j = v_i^T (\lambda_j v_j) = 0$$

então $\{A v_1, \dots, A v_n\}$ é um conjunto ortogonal. Considerando que os comprimentos dos vetores $A v_1, \dots, A v_n$ são dados pelos valores singulares de A e que existem r valores sin-

gulares não-nulos, sendo $Av_i \neq 0$ se e somente se $1 \leq i \leq r$, então Av_1, \dots, Av_r são vetores linearmente independentes e pertencem ao espaço coluna de A .

Por último, qualquer y pertencente ao $\mathcal{R}(A)$, por exemplo $y = Ax$, pode ser escrito como

$$y = Ax = c_1 Av_1 + \dots + c_r Av_r + c_{r+1} Av_{r+1} + \dots + c_n Av_n$$

ou

$$y = c_1 Av_1 + \dots + c_r Av_r + 0 + \dots + 0$$

portanto y pertence ao $\text{Span}\{Av_1, \dots, Av_r\}$, mostrando assim que $\{Av_1, \dots, Av_r\}$ é uma base ortogonal para $\mathcal{R}(A)$ e conseqüentemente que $\text{posto de } A = \dim \mathcal{R}(A) = r$. A forma mais confiável computacionalmente de se estimar o posto de uma matriz de grandes dimensões é contar o número de valores singulares não-nulos, neste caso, se o valor singular for extremamente pequeno, este é considerado nulo.

TEOREMA : Seja $A_{m \times n}$ uma matriz de posto r . Então existe uma matriz $\Sigma_{m \times n} = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$ tal que os elementos diagonais de $D_{r \times r}$ são os r primeiros valores singulares de A , $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$, e existem matrizes ortogonais $U_{m \times m}$ e $V_{n \times n}$, tais que $A = U \Sigma V^T$.

DEMONSTRAÇÃO [LAY97]: Sejam λ_i e v_i como no teorema anterior, de tal forma que $\{Av_1, \dots, Av_r\}$ seja uma base ortogonal para o espaço coluna de A , $\mathcal{R}(A)$. Normalizando cada Av_i obtemos uma base ortonormal $\{u_1, \dots, u_r\}$, como adiante

$$u_i = \frac{1}{\|Av_i\|} Av_i = \frac{1}{\sigma_i} Av_i$$

que resulta

$$Av_i = u_i \sigma_i, \quad (i = 1, \dots, r).$$

Estendendo para uma base ortonormal $\{u_1, \dots, u_m\}$ de \mathbb{R}^m e considerando as matrizes ortogonais $U = (u_1 \ u_2 \ \dots \ u_m)$ e $V = (v_1 \ v_2 \ \dots \ v_n)$, tem-se

$$AV = (Av_1 \ \dots \ Av_r \ 0 \ \dots \ 0) = (u_1\sigma_1 \ \dots \ u_r\sigma_r \ 0 \ \dots \ 0).$$

Considerando D e Σ como define-se anteriormente, obtém-se

$$U\Sigma = (u_1 \ u_2 \ \dots \ u_m) \begin{pmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_r & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

ou

$$U\Sigma = (u_1\sigma_1 \ \dots \ u_r\sigma_r \ 0 \ \dots \ 0) = AV.$$

E então sendo V uma matriz ortogonal, obtém-se

$$U\Sigma V^T = AVV^T = A \tag{19}$$

Observa-se que as matrizes U e V não são unicamente determinadas por A , porém os r primeiros elementos diagonais de Σ são seus valores singulares.

2.1.9 Decomposição de Valor Singular Reduzida

Se Σ possuir linhas ou colunas nulas, então pode-se obter uma decomposição mais compacta para A . Considerando a notação estabelecida anteriormente e particionando as

matrizes $U = (u_1 \cdots u_r \cdots u_m)$ e $V = (v_1 \cdots v_r \cdots v_n)$, obtém-se

$$\begin{aligned}
 A &= U\Sigma V^T \\
 A &= (U_r \quad U_{m-r}) \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_r^T \\ V_{n-r}^T \end{pmatrix} \\
 A &= U_r D V_r^T.
 \end{aligned} \tag{20}$$

Essa fatoração é denominada *decomposição de valor singular reduzida*.

2.1.10 Inversa Generalizada ou Pseudo-Inversa

Suponha que a matriz $A_{m \times n}$ de posto r possua uma decomposição de valor singular $A = U\Sigma V^T$, com $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. Então, representa-se por A^+ e denomina-se de *inversa generalizada* ou *pseudo inversa de Moore-Penrose* a matriz $n \times m$, determinada por

$$A^+ = V\Sigma^+U^T, \tag{21}$$

onde $\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$ como definido anteriormente e $\Sigma^+ = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix}$. Nestas condições, observa-se que $D_{r \times r}^{-1}$ é uma matriz diagonal, cujos elementos diagonais são os inversos dos r primeiros valores singulares de A , $0 < \sigma_1^{-1} \leq \sigma_2^{-1} \leq \cdots \leq \sigma_r^{-1}$.

PROPRIEDADES :

- (i) $AA^+A = A$,
- (ii) $A^+AA^+ = A^+$ e
- (iii) AA^+ e A^+A são hermitianas.

Além disso, a inversa generalizada A^+ é unicamente determinada por A , embora U e V , usados na decomposição de valores singulares, não sejam necessariamente únicos.

Agora, com a definição da *pseudo-inversa*, pode-se generalizar o número de condição $c = \|A\| \|A^{-1}\|$, definido anteriormente. Para o caso de A ser uma matriz retangular, tem-se

$$c = \|A\| \|A^+\|. \quad (22)$$

Considerando os r primeiros valores singulares de A ordenados de tal forma que $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$, obtém-se conseqüentemente os valores singulares de A^+ , seguindo $0 < \sigma_1^{-1} \leq \sigma_2^{-1} \leq \dots \leq \sigma_r^{-1}$. Assim, levando em consideração os conhecimentos já expostos neste trabalho, pode-se escrever

$$\|A\| = [\lambda_{\max} \text{ de } A^T A]^{1/2} = \sigma_{\max} \text{ de } A, \quad (23)$$

$$\|A^+\| = [\lambda_{\min} \text{ de } A^T A]^{-1/2} = \sigma_{\min}^{-1} \text{ de } A \text{ e} \quad (24)$$

$$c = \|A\| \|A^+\| = \left(\frac{\lambda_{\max}}{\lambda_{\min}} \right)^{1/2} = \frac{\sigma_{\max}}{\sigma_{\min}}. \quad (25)$$

2.1.11 Formas Quadráticas e Autovalores

Denomina-se de forma quadrática em \mathbb{R}^n uma função Q de \mathbb{R}^n em \mathbb{R}^n , definida por $Q(x) = x^T A x$, onde x é um vetor de \mathbb{R}^n e $A_{n \times n}$ é uma matriz simétrica. Considerando a_{ij} a representação dos elementos de A , pode-se também escrever $Q(x) = \sum_{i=1}^n \sum_{j=1}^m a_{ij} x_i x_j$. É possível distinguir várias classes de formas quadráticas pelo tipo de valores que estas assumem quando x varia em \mathbb{R}^n :

Positiva definida se $Q(x) > 0$ para todo $x \neq 0$,

Negativa definida se $Q(x) < 0$ para todo $x \neq 0$ e

Indefinida se $Q(x)$ adquire valores tanto positivos quanto negativos.

Além disso, Q é denominada *positiva semidefinida* se $Q(x) \geq 0$ para todo x , e *negativa semidefinida* se $Q(x) \leq 0$ para todo x .

TEOREMA : Seja A uma matriz real simétrica $n \times n$. Então, A é positiva definida se e somente se todos os seus autovalores são positivos .

DEMONSTRAÇÃO [STR93]: Se A é positiva definida e se λ é um autovalor de A , então, para qualquer autovetor de A associado a λ , tem-se $x^T Ax = \lambda x^T x = \lambda \|x\|^2$, que produz

$$\lambda = \frac{x^T Ax}{\|x\|^2} > 0.$$

Reciprocamente, supondo agora todos os autovalores de A positivos e um conjunto ortonormal de autovetores associados $\{x_1, \dots, x_n\}$. Sendo x um vetor não nulo em \mathbb{R}^n , então x pode ser escrito como

$$x = \alpha_1 x_1 + \dots + \alpha_n x_n$$

onde $\alpha_i = x^T x_i$ para $i = 1, 2, \dots, n$ e $\sum_{i=1}^n (\alpha_i)^2 = \|x\|^2 > 0$. Então,

$$x^T Ax = (\alpha_1 x_1 + \dots + \alpha_n x_n)^T (\alpha_1 \lambda_1 x_1 + \dots + \alpha_n \lambda_n x_n)$$

$$x^T Ax = \sum_{i=1}^n (\alpha_i)^2 \lambda_i$$

logo,

$$x^T Ax \geq (\min \lambda_i) \|x\|^2 > 0 \tag{26}$$

e, portanto, A é *positiva definida*. Se todos autovalores da *matriz simétrica* A forem negativos, então $-A$ será positiva definida e por conseguinte A será *negativa definida*. Se os

autovalores de A têm sinais diferentes, então A é *indefinida*, pois se λ_1 for positivo com o autovetor associado x_1 , tem-se $x_1^T A x_1 = \lambda_1 x_1^T x_1 = \lambda_1 \|x_1\|^2 > 0$, em contrapartida se, por exemplo, λ_2 for negativo, então $x_2^T A x_2 = \lambda_2 x_2^T x_2 = \lambda_2 \|x_2\|^2 < 0$.

2.1.12 Raiz Quadrada de Uma Matriz

Seja Σ uma matriz $p \times p$ *definida positiva* com a decomposição espectral

$$\Sigma = \sum_{i=1}^p \lambda_i e_i e_i^T$$

e a matriz dos autovetores normalizados $P_{p \times p} = (e_1 \ e_2 \ \dots \ e_p)$, então tem-se

$$\Sigma = \sum_{i=1}^p \lambda_i e_i e_i^T = P \Lambda P^T, \quad (27)$$

onde $P^T P = P P^T = I$ e $\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \end{pmatrix}$ com $\lambda_i \geq 0$.

Assim,

$$\Sigma^{-1} = P \Lambda^{-1} P^T = \sum_{i=1}^p \frac{1}{\lambda_i} e_i e_i^T, \quad (28)$$

desde que

$$(P \Lambda P^T) (P \Lambda^{-1} P^T) = (P \Lambda^{-1} P^T) (P \Lambda P^T) = P P^T = I.$$

Considerando ainda que

$$\Lambda^{1/2} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{\lambda_p} \end{pmatrix}_{p \times p},$$

então,

$$\Sigma^{1/2} = \sum_{i=1}^p \sqrt{\lambda_i} e_i e_i^T = P \Lambda^{1/2} P^T. \quad (29)$$

PROPRIEDADES :

$$(i) \quad (\Sigma^{1/2})^T = \Sigma^{1/2} \quad (\Sigma^{1/2} \text{ é simétrica}),$$

$$(ii) \quad \Sigma^{1/2} \Sigma^{1/2} = \Sigma,$$

$$(iii) \quad (\Sigma^{1/2})^{-1} = \sum_{i=1}^p \frac{1}{\sqrt{\lambda_i}} e_i e_i^T = P \Lambda^{-1/2} P^T \text{ onde } \Lambda^{-1/2} \text{ é diagonal com o } i - \\ \text{ésimo elemento } \frac{1}{\sqrt{\lambda_i}},$$

$$(iv) \quad \Sigma^{1/2} \Sigma^{-1/2} = \Sigma^{-1/2} \Sigma^{1/2} = I \text{ e } \Sigma^{-1/2} \Sigma^{-1/2} = \Sigma^{-1} \text{ onde } \Sigma^{-1/2} = (\Sigma^{1/2})^{-1}.$$

2.2 Análise de Correlação Canônica

2.2.1 Vetores e Matrizes Aleatórias

São denominados matriz aleatória ou vetor aleatório, a matriz ou vetor cujos elementos são variáveis aleatórias. O valor esperado é dado pela matriz ou vetor cujos elementos são os valores esperados de cada variável aleatória. Especificamente para uma matriz aleatória $X = \{X_{ij}\}_{n \times p}$, tem-se

$$E(X) = \begin{pmatrix} E(X_{11}) & E(X_{12}) & \cdots & E(X_{1p}) \\ E(X_{21}) & E(X_{22}) & \cdots & E(X_{2p}) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_{n1}) & E(X_{n2}) & \cdots & E(X_{np}) \end{pmatrix},$$

onde

$$E(X_{ij}) = \begin{cases} \int_{-\infty}^{\infty} x_i f_i(x_i) dx_i, & \text{se } X_i \text{ é variável aleatória contínua} \\ \sum_{\text{todo } x_i} x_i g_i(x_i), & \text{se } X_i \text{ é variável aleatória discreta} \end{cases}$$

com função de densidade de probabilidade $f_i(x_i)$.
com função de probabilidade $g_i(x_i)$.

Dois resultados importantes, envolvendo a soma e produto de matrizes, são obtidos diretamente da definição do valor esperado de uma matriz aleatória e das propriedades da esperança de uma variável,

$$E(X_1 + Y_1) = E(X_1) + E(Y_1) \quad \text{e} \quad E(cX_1) = cE(X_1).$$

Sejam X e Y duas matrizes de variáveis aleatórias de mesma dimensão, e A e B matrizes conformes, constituídas por constantes. Então, valem as propriedades:

$$E(X + Y) = E(X) + E(Y)$$

$$E(AXB) = AE(X)B \tag{30}$$

2.2.2 Vetor de Médias, Matriz de Covariâncias e Matriz de Correlação.

Supondo $X = [X_1, X_2, \dots, X_p]^T$ um vetor aleatório e que cada um de seus elementos X_i ($i = 1, 2, \dots, p$) seja uma variável aleatória com sua própria distribuição marginal de probabilidade. As médias marginais μ_i e variâncias σ_{ii} são definidas como

$$\mu_i = E(X_i)$$

e

$$\sigma_{ii} = E(X_i - \mu_i)^2,$$

onde

$$\mu_i = \begin{cases} \int_{-\infty}^{\infty} x_i f_i(x_i) dx_i, & \text{se } X_i \text{ é uma variável aleatória contínua} \\ \sum_{\forall x_i} x_i g_i(x_i), & \text{se } X_i \text{ é variável aleatória discreta} \end{cases}$$

com função de densidade de probabilidade $f_i(x_i)$.
com função de probabilidade $g_i(x_i)$.

e

$$\sigma_i^2 = \begin{cases} \int_{-\infty}^{\infty} (x_i - \mu_i)^2 f_i(x_i) dx_i, & \text{se } X_i \text{ é uma variável aleatória contínua} \\ \sum_{\forall x_i} (x_i - \mu_i)^2 g_i(x_i), & \text{se } X_i \text{ é variável aleatória discreta} \end{cases}$$

com função de densidade de probabilidade $f_i(x_i)$.
com função de probabilidade $g_i(x_i)$.

(31)

O comportamento de algum par de variáveis (X_i, X_k) , componentes do vetor X , é descrito por sua função de probabilidade conjunta e a *medida da associação linear* entre elas é dada pela covariância,

$$\sigma_{ik} = E[(X_i - \mu_i)(X_k - \mu_k)] \quad (32)$$

ou seja,

$$\sigma_{i k} = \begin{cases} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - \mu_i) (x_k - \mu_k) f_{i k} (x_i, x_k) dx_i dx_k, & \text{se } X_i \text{ e } X_k \text{ são variáveis} \\ & \text{aleatórias contínuas com função de densidade de probabilidade } f_{i k} (x_i, x_k), \\ \sum_{\forall x_i} \sum_{\forall x_k} (x_i - \mu_i) (x_k - \mu_k) g_{i k} (x_i, x_k), & \text{se } X_i \text{ e } X_k \text{ são variáveis} \\ & \text{aleatórias discretas com função de probabilidade } g_{i k} (x_i, x_k), \end{cases} \quad (33)$$

sendo μ_i e μ_k ($i, k = 1, 2, \dots, p$), as médias marginais. Quando $i = k$, a covariância torna-se a variância marginal.

Generalizando o comportamento coletivo das variáveis aleatórias X_1, X_2, \dots, X_p ou, e-
quivalentemente, do vetor $X = (X_1 \ X_2 \ \dots \ X_p)^T$ é descrito pela função de densi-
dade de probabilidade conjunta $f(x) = (x_1, x_2, \dots, x_p)$.

Na prática, $f(x)$ é freqüentemente uma função multivariada de densidade normal. Se $P[X_i \leq x_i \text{ e } X_k \leq x_k]$, probabilidade conjunta, pode ser escrita como produto das cor-
respondentes probabilidades marginais, tal que,

$$P[X_i \leq x_i \text{ e } X_k \leq x_k] = P[X_i \leq x_i] P[X_k \leq x_k]$$

para todos os pares de valores x_i, x_k , então X_i e X_k são ditas estatisticamente *independen-*
tes. Quando X_i e X_k são variáveis aleatórias contínuas com densidade conjunta $f_{i k}(x_i, x_k)$
e densidades marginais $f_i(x_i)$ e $f_k(x_k)$, a condição de independência torna-se

$$f_{i k}(x_i, x_k) = f_i(x_i) f_k(x_k)$$

para todos os pares (x_i, x_k) . As p variáveis aleatórias contínuas X_1, X_2, \dots, X_p são, estatisti-
camente, ditas *mutuamente independentes* se a densidade conjunta pode ser fatorada como

$$f_{12\dots p}(x_1, x_2, \dots, x_p) = f_1(x_1) f_2(x_2) \dots f_p(x_p)$$

para todas p -nuplas (x_1, x_2, \dots, x_p) . A independência estatística tem uma implicação importante para a covariância, pois a fatoração anterior implica que a covariância de (X_i, X_k) seja nula. Então, $Cov(X_i, X_k) = 0$ se X_i e X_k são independentes. A recíproca, em geral, não é verdadeira pois existem situações em que $Cov(X_i, X_k) = 0$, mas X_i e X_k não são independentes. Assim, pode-se definir o *coeficiente de correlação linear*, entre um par de variáveis (X_i, X_k) , como sendo o quociente ρ_{ik} , dado por

$$\rho_{ik} = \frac{\sigma_{ik}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{kk}}},$$

onde

$$\sigma_{ik} = Cov(X_i, X_k) = E(X_i - \mu_i)(X_k - \mu_k), \quad (34)$$

$$\sigma_{ii} = Var(X_i) = E(X_i - \mu_i)^2$$

e

$$\sigma_{kk} = Var(X_k) = E(X_k - \mu_k)^2. \quad (35)$$

2.3 Correlação Canônica

2.3.1 Introdução

A análise de correlação canônica tem como objetivo identificar e quantificar as associações lineares entre *dois conjuntos* de variáveis aleatórias. No desenvolvimento dessa análise, determina-se primeiramente o par de combinações lineares, uma de cada conjunto de variáveis, com maior correlação; em seguida, determina-se um segundo par de combinações lineares de maior correlação entre todos os pares não correlacionados com aqueles selecionados inicialmente, e assim por diante até o $p - \text{ésimo}$ elemento sendo $p \times p$ a ordem da matriz de correlação. Uma aplicação dessa análise ocorre em processos de produção contínua, quando se-deseja prever o comportamento de um conjunto de variáveis a partir do conhecimento prévio de um outro conjunto de variáveis. Considerando-se um desses conjuntos como um conjunto de variáveis independentes, e o outro como um conjunto de variáveis dependentes, e, se existe uma relação linear entre esses conjuntos, então pode-se calcular a magnitude dessa relação, a qual denominamos *correlação canônica*.

2.3.2 Variáveis Canônicas e Correlações Canônicas

Para o desenvolvimento deste tema vai-se admitir, como exemplo, o estudo das medidas de associação entre dois conjuntos de variáveis aleatórias, cada qual com sua própria distribuição de probabilidade marginal. O primeiro conjunto é denominado de X , composto de variáveis aleatórias X_i , ($i = 1, 2, \dots, p$), e o segundo de Y , composto de variáveis aleatórias Y_j , ($j = 1, 2, \dots, q$). Assumindo $p \leq q$ e representando os conjuntos como vetores

aleatórios tem-se:

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix} \quad \text{e} \quad Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_q \end{pmatrix} .$$

Assim, os vetores dos valores esperados das variáveis aleatórias serão representados por

$$E(X) = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_p) \end{pmatrix} \quad \text{e} \quad E(Y) = \begin{pmatrix} E(Y_1) \\ \vdots \\ E(Y_q) \end{pmatrix} ,$$

onde para cada elemento, tem-se

$$E(X_i) = \begin{cases} \int_{-\infty}^{\infty} x_i f_i(x_i) dx_i , & \text{se } X_i \text{ é variável aleatória contínua} \\ \sum_{\forall x_i} x_i g_i(x_i) , & \text{se } X_i \text{ é variável aleatória discreta} \end{cases}$$

com função de densidade de probabilidade $f_i(x_i)$,
com função de probabilidade $g_i(x_i)$,

e o equivalente para $E(Y_j)$. As médias marginais μ_{x_i} e μ_{y_j} são dadas por $\mu_{x_i} = E(X_i)$ e

$\mu_{y_j} = E(Y_j)$ e os vetores das médias marginais são então representados por

$$\mu_x = E(X) = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_p) \end{pmatrix} = \begin{pmatrix} \mu_{x1} \\ \vdots \\ \mu_{xp} \end{pmatrix}$$

e

$$\mu_y = E(Y) = \begin{pmatrix} E(Y_1) \\ \vdots \\ E(Y_q) \end{pmatrix} = \begin{pmatrix} \mu_{y1} \\ \vdots \\ \mu_{yq} \end{pmatrix} .$$

Como essa análise recai sobre dois conjuntos de variáveis aleatórias X e Y , para determinar as medidas de dispersão de todas as interrelações entre as variáveis dos dois conjuntos, é necessário considerar conjuntamente os vetores X e Y , em um único vetor Z , como segue,

$$Z = \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \\ Y_1 \\ \vdots \\ Y_q \end{pmatrix} = \begin{pmatrix} Z_1 \\ \vdots \\ Z_p \\ Z_{p+1} \\ \vdots \\ Z_{p+q} \end{pmatrix} .$$

Logo o vetor das médias marginais fica

$$\mu = E(Z) = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} = \begin{pmatrix} E(X) \\ E(Y) \end{pmatrix}, \quad (36)$$

ou seja,

$$\mu = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_p) \\ E(Y_1) \\ \vdots \\ E(Y_q) \end{pmatrix} = \begin{pmatrix} \mu_{x1} \\ \vdots \\ \mu_{xp} \\ \mu_{y1} \\ \vdots \\ \mu_{yq} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \\ \mu_{p+1} \\ \vdots \\ \mu_{p+q} \end{pmatrix}.$$

Como já vimos, a matriz das covariâncias das variáveis componentes de Z será dada por

$$COV(Z) = E \left[(Z - \mu)(Z - \mu)^T \right], \quad (37)$$

ou seja,

$$E \left[\begin{pmatrix} X_1 - \mu_{x1} \\ \vdots \\ X_p - \mu_{xp} \\ Y_1 - \mu_{y1} \\ \vdots \\ Y_q - \mu_{yq} \end{pmatrix} \begin{pmatrix} X_1 - \mu_{x1} & \cdots & X_p - \mu_{xp} & Y_1 - \mu_{y1} & \cdots & Y_q - \mu_{yq} \end{pmatrix} \right]$$

que produz

$$E \begin{pmatrix} (X_1 - \mu_{x1})^2 & \cdots & (X_1 - \mu_{x1})(X_p - \mu_{xp}) & (X_1 - \mu_{x1})(Y_1 - \mu_{y1}) & \cdots & (X_1 - \mu_{x1})(Y_q - \mu_{yq}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (X_p - \mu_{xp})(X_1 - \mu_{x1}) & \cdots & (X_p - \mu_{xp})^2 & (X_p - \mu_{xp})(Y_1 - \mu_{y1}) & \cdots & (X_p - \mu_{xp})(Y_q - \mu_{yq}) \\ (Y_1 - \mu_{y1})(X_1 - \mu_{x1}) & \cdots & (Y_1 - \mu_{y1})(X_p - \mu_{xp}) & (Y_1 - \mu_{y1})^2 & \cdots & (Y_1 - \mu_{y1})(Y_q - \mu_{yq}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (Y_q - \mu_{yq})(X_1 - \mu_{x1}) & \cdots & (Y_q - \mu_{yq})(X_p - \mu_{xp}) & (Y_q - \mu_{yq})(Y_1 - \mu_{y1}) & \cdots & (Y_q - \mu_{yq})^2 \end{pmatrix}$$

ou, equivalentemente,

$$E \begin{pmatrix} (Z_1 - \mu_1)^2 & \cdots & (Z_1 - \mu_1)(Z_p - \mu_p) & (Z_1 - \mu_1)(Z_{p+1} - \mu_{p+1}) & \cdots & (Z_1 - \mu_1)(Z_{p+q} - \mu_{p+q}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (Z_p - \mu_p)(Z_1 - \mu_1) & \cdots & (Z_p - \mu_p)^2 & (Z_p - \mu_p)(Z_{p+1} - \mu_{p+1}) & \cdots & (Z_p - \mu_p)(Z_{p+q} - \mu_{p+q}) \\ (Z_{p+1} - \mu_{p+1})(Z_1 - \mu_1) & \cdots & (Z_{p+1} - \mu_{p+1})(Z_p - \mu_p) & (Z_{p+1} - \mu_{p+1})^2 & \cdots & (Z_{p+1} - \mu_{p+1})(Z_{p+q} - \mu_{p+q}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (Z_{p+q} - \mu_{p+q})(Z_1 - \mu_1) & \cdots & (Z_{p+q} - \mu_{p+q})(Z_p - \mu_p) & (Z_{p+q} - \mu_{p+q})(Z_{p+1} - \mu_{p+1}) & \cdots & (Z_{p+q} - \mu_{p+q})^2 \end{pmatrix}$$

Considerando $\sigma_{IK} = E[(Z_I - \mu_I)(Z_K - \mu_K)]$ e $\sigma_{II} = E(Z_I - \mu_I)^2$, obtém-se

$$COV(Z) = \Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1p} & \sigma_{1,p+1} & \cdots & \sigma_{1,p+q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} & \sigma_{p,p+1} & \cdots & \sigma_{p,p+q} \\ \sigma_{p+1,1} & \cdots & \sigma_{p+1,p} & \sigma_{p+1,p+1} & \cdots & \sigma_{p+1,p+q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{p+q,1} & \cdots & \sigma_{p+q,p} & \sigma_{p+q,p+1} & \cdots & \sigma_{p+q,p+q} \end{pmatrix}$$

Como $\sigma_{IK} = E[(Z_I - \mu_I)(Z_K - \mu_K)]$ é igual a $\sigma_{KI} = E[(Z_K - \mu_K)(Z_I - \mu_I)]$, para $(I, K = 1, 2, \dots, (p+q))$, então Σ é uma matriz simétrica e pode-se escrever

$$COV(Z) = \Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1p} & \sigma_{1,p+1} & \cdots & \sigma_{1,p+q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \cdots & \sigma_{pp} & \sigma_{p,p+1} & \cdots & \sigma_{p,p+q} \\ \sigma_{1,p+1} & \cdots & \sigma_{p,p+1} & \sigma_{p+1,p+1} & \cdots & \sigma_{p+1,p+q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{1,p+q} & \cdots & \sigma_{p,p+q} & \sigma_{p+1,p+q} & \cdots & \sigma_{p+q,p+q} \end{pmatrix}.$$

Como já visto, o coeficiente de correlação ρ_{IK} é definido em função da covariância σ_{IK} e das variâncias σ_{II} e σ_{KK} , por

$$\rho_{IK} = \frac{\sigma_{IK}}{\sqrt{\sigma_{II}\sigma_{KK}}}$$

O coeficiente de correlação mede a magnitude da associação linear entre as variáveis aleatórias componentes do vetor Z e por consequência entre as variáveis do vetor X , entre as variáveis do vetor Y e entre as variáveis do vetor X e Y . Considerando-se a matriz ρ de coeficientes de correlação dada por

$$\rho = \begin{pmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{11}}} & \cdots & \frac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} & \frac{\sigma_{1p+1}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{p+1p+1}}} & \cdots & \frac{\sigma_{1p+q}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{p+qp+q}}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} & \cdots & \frac{\sigma_{pp}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{pp}}} & \frac{\sigma_{pp+1}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{p+1p+1}}} & \cdots & \frac{\sigma_{pp+q}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{p+qp+q}}} \\ \frac{\sigma_{1p+1}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{p+1p+1}}} & \cdots & \frac{\sigma_{pp+1}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{p+1p+1}}} & \frac{\sigma_{p+1p+1}}{\sqrt{\sigma_{p+1p+1}}\sqrt{\sigma_{p+1p+1}}} & \cdots & \frac{\sigma_{p+1p+q}}{\sqrt{\sigma_{p+1p+1}}\sqrt{\sigma_{p+qp+q}}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{1p+q}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{p+qp+q}}} & \cdots & \frac{\sigma_{pp+q}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{p+qp+q}}} & \frac{\sigma_{p+1p+q}}{\sqrt{\sigma_{p+1p+1}}\sqrt{\sigma_{p+qp+q}}} & \cdots & \frac{\sigma_{p+qp+q}}{\sqrt{\sigma_{p+qp+q}}\sqrt{\sigma_{p+qp+q}}} \end{pmatrix}$$

ou

$$\rho = \begin{pmatrix} 1 & \cdots & \rho_{1p} & \rho_{1p+1} & \cdots & \rho_{1p+q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \rho_{1p} & \cdots & 1 & \rho_{pp+1} & \cdots & \rho_{pp+q} \\ \rho_{1p+1} & \cdots & \rho_{pp+1} & 1 & \cdots & \rho_{p+1p+q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \rho_{1p+q} & \cdots & \rho_{pp+q} & \rho_{p+1p+q} & \cdots & 1 \end{pmatrix} = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix}$$

e a matriz $D^{1/2}$ dos desvios padrão

$$D^{1/2} = \begin{pmatrix} \sqrt{\sigma_{11}} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\sigma_{pp}} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \sqrt{\sigma_{p+1p+1}} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \sqrt{\sigma_{p+qp+q}} \end{pmatrix},$$

então, é fácil verificar que

$$D^{1/2}\rho D^{1/2} = \Sigma$$

e

$$\rho = (D^{1/2})^{-1} \Sigma (D^{1/2})^{-1}.$$

Isto é, Σ pode ser obtida de $D^{1/2}$ e ρ , assim como, ρ pode ser obtido de Σ . Reconsiderando a matriz de covariâncias Σ particionada por quatro blocos, conforme a seguir

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

observa-se que:

- (i) As matrizes Σ_{11} e Σ_{22} são simétricas.
- (ii) As matrizes Σ_{12} e Σ_{21} não são necessariamente simétricas ou quadradas.
- (iii) A matriz Σ_{12} é igual transposta da matriz Σ_{21} , ou seja, $\Sigma_{12} = \Sigma_{21}^T$.
- (iv) As variâncias-covariâncias entre os pares de variáveis componentes do vetor (conjunto) X estão contidas na matriz Σ_{11} , ou seja, $Cov(X) = \Sigma_{11}$.
- (v) As variâncias-covariâncias entre os pares de variáveis componentes do vetor (conjunto) Y estão contidas na matriz Σ_{22} , ou seja, $Cov(Y) = \Sigma_{22}$.
- (vi) As covariâncias entre os pares de variáveis de diferentes vetores (conjuntos), uma variável de X e uma variável de Y estão contidas na matriz de “covariância cruzada” Σ_{12} ou, equivalentemente, em Σ_{21} , ou seja, $Cov(X, Y) = \Sigma_{12}$ ou $Cov(X, Y) = \Sigma_{21}$.

Então, os elementos de Σ_{12} ou Σ_{21} medem a associação entre os dois conjuntos. Quando p e q são relativamente grandes, a interpretação coletiva dos elementos de Σ_{12} ou Σ_{21} é normalmente difícil de ser feita. Isto ocorre frequentemente nas combinações lineares de variáveis que são de interesse e úteis para finalidade de previsão e comparação. A principal tarefa da análise de correlação canônica é resumir as associações entre

os conjuntos de variáveis X e Y em termos de uma escolha cuidadosa de poucas covariâncias (ou correlações) preferencialmente as pq covariâncias de Σ_{12} . As combinações lineares fornecem um resumo das medidas de um conjunto de variáveis.

Sejam $U = a^T X$ e $V = b^T Y$ para algum par de vetores de coeficientes a e b .

Então tem-se

$$U = a^T X \Rightarrow U = (a_1 \ a_2 \ \cdots \ a_p) \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix}$$

$$U = a_1 X_1 + a_2 X_2 + \cdots + a_p X_p$$

$$E(U) = E(a_1 X_1 + a_2 X_2 + \cdots + a_p X_p)$$

$$E(U) = (a_1 E(X_1) + a_2 E(X_2) + \cdots + a_p E(X_p)) = (a_1 \mu_{x1} + a_2 \mu_{x2} + \cdots + a_p \mu_{xp})$$

$$E(U) = (a_1 \ a_2 \ \cdots \ a_p) \begin{pmatrix} \mu_{x1} \\ \mu_{x2} \\ \vdots \\ \mu_{xp} \end{pmatrix} = a^T \mu_x$$

$$Var(U) = Var(a_1 X_1 + a_2 X_2 + \cdots + a_p X_p)$$

$$Var(U) = a^T Var(X) a$$

$$Var(U) = a^T \Sigma_{11} a \quad (38)$$

Equivalentemente, demonstra-se ainda que

$$Var(V) = b^T \Sigma_{22} b$$

e

$$Cov(U, V) = a^T \Sigma_{12} b. \quad (39)$$

Devemos procurar os vetores de coeficientes a e b tal que a correlação

$$Corr(U, V) = \frac{Cov(U, V)}{\sqrt{Var(U)} \sqrt{Var(V)}} \quad (40)$$

ou

$$\text{Corr}(U, V) = \frac{a^T \Sigma_{12} b}{\sqrt{a^T \Sigma_{11} a} \sqrt{b^T \Sigma_{22} b}} \quad (41)$$

seja a maior possível.

Serão seguidas as seguintes definições:

- (i) O primeiro par de variáveis canônicas é o par da combinação linear U_1, V_1 , com variâncias unitárias e que maximiza a correlação $\text{Corr}(U, V)$;
- (ii) O segundo par de variáveis canônicas é o par da combinação linear U_2, V_2 , com variâncias unitárias, o qual maximiza a correlação $\text{Corr}(U, V)$, entre todas as escolhas não correlacionadas com o primeiro par de variáveis canônicas;
- (iii) O K -ésimo par de variáveis canônicas é o par da combinação linear U_k, V_k , com variâncias unitárias, o qual maximiza a relação $\text{Corr}(U, V)$, entre todas as escolhas não correlacionadas com os $K - 1$ pares de variáveis canônicas anteriores .

A correlação entre $K - \text{ésimo}$ par de variáveis canônicas é chamada de $K - \text{ésima}$ correlação canônica.

A demonstração do teorema a seguir dá o detalhamento necessário para a obtenção das variáveis canônicas e suas correlações.

TEOREMA: Supondo os vetores aleatórios $X_{(p \times 1)}$ e $Y_{(q \times 1)}$ conforme já definidos anteriormente com $p \geq q$, $\text{Cov}(X) = \Sigma_{11}$, $\text{Cov}(Y) = \Sigma_{22}$, e $\text{Cov}(X, Y) = \Sigma_{12}$. Supondo ainda que a matriz $\text{Cov}(Z) = \Sigma$ tem *posto completo*. Sejam as combinações lineares dadas por $U = a^T X$ e $V = b^T Y$, onde $a_{(p \times 1)}$ e $b_{(q \times 1)}$ são os vetores de coeficientes, então,

$$\max_{a,b} \text{Corr}(U, V) = \rho_1^* \quad (42)$$

é atingida pelo primeiro par de variáveis canônicas,

$$U_1 = a_1^T X \quad \text{e} \quad V_1 = b_1^T Y \quad (43)$$

ou

$$U_1 = e_1^T \Sigma_{11}^{-1/2} X \quad \text{e} \quad V_1 = f_1^T \Sigma_{22}^{-1/2} Y. \quad (44)$$

O k -ésimo par de variáveis canônicas, $k = 2, 3, \dots, p$, dado por

$$U_k = e_k^T \Sigma_{11}^{-1/2} X \quad \text{e} \quad V_k = f_k^T \Sigma_{22}^{-1/2} Y,$$

maximiza $\text{Corr}(U_k, V_k) = \rho_k^*$, entre aquelas combinações lineares não correlacionadas com os $K - 1$ pares de variáveis canônicas anteriores. Aqui $\rho_1^{*2} \geq \rho_2^{*2} \geq \dots \geq \rho_p^{*2}$ são os autovalores de $\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2}$ e e_1, e_2, \dots, e_p são autovetores associados (Os valores $\rho_1^{*2}, \rho_2^{*2}, \dots, \rho_p^{*2}$ são também os p maiores autovalores de $\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1/2}$ com os correspondentes autovetores f_1, f_2, \dots, f_p sendo cada f_i proporcional a $\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1} e_i$.

As variáveis canônicas têm as seguintes propriedades:

- (i) $\text{Var}(U_k) = \text{Var}(V_k) = 1$;
- (ii) $\text{Cov}(U_k, U_L) = \text{Corr}(U_k, U_L) = 0, K \neq L$;
- (iii) $\text{Cov}(V_k, V_L) = \text{Corr}(V_k, V_L) = 0, K \neq L$;
- (iv) $\text{Cov}(U_k, V_L) = \text{Corr}(U_k, V_L) = 0, K \neq L$

para $K, L = 1, 2, 3, \dots, p$.

DEMONSTRAÇÃO[JON88]: Supondo Σ_{11} e Σ_{12} matrizes *simétricas e não singulares*, e usando o conceito de raiz quadrada de uma matriz, obtém-se:

$$\Sigma_{11} = \Sigma_{11}^{1/2} \Sigma_{11}^{1/2}$$

$$\Sigma_{11}^{-1} = \Sigma_{11}^{-1/2} \Sigma_{11}^{-1/2}$$

$$c = \Sigma_{11}^{1/2} a$$

$$a = \Sigma_{11}^{-1/2} c,$$

e similarmente

$$b = \Sigma_{22}^{-1/2} d . \quad (45)$$

Deve-se, então procurar a e b que torne a correlação

$$\text{Corr}(U, V) = \frac{a^T \Sigma_{12} b}{\sqrt{a^T \Sigma_{11} a} \sqrt{b^T \Sigma_{22} b}} \quad (46)$$

a maior possível, onde $U = a^T X$ e $V = b^T Y$. Substituindo ?? e 2.45 em 2.46, obtém-se

$$\text{Corr}(U, V) = \frac{c^T \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} d}{\sqrt{c^T c} \sqrt{d^T d}} . \quad (47)$$

Da desigualdade de Cauchy-Schwartz,

$$(\alpha^T \beta)^2 \leq (\alpha^T \alpha) (\beta^T \beta)$$

obtém-se

$$(\alpha^T \beta) \leq (\alpha^T \alpha)^{1/2} (\beta^T \beta)^{1/2} .$$

Observando que a igualdade ocorre quando α é proporcional a β ou vice-versa e aplicando essa desigualdade ao numerador em 2.47, obtém-se

$$\underbrace{c^T \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2}}_{\alpha^T} \underbrace{d}_{\beta} \leq \left(\underbrace{c^T \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2}}_{\alpha^T} \underbrace{\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} c}_{\alpha} \right)^{1/2} \left(\underbrace{d^T}_{\beta^T} \underbrace{d}_{\beta} \right)^{1/2}$$

que produz

$$c^T \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} d \leq \left(c^T \underbrace{\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2}} c \right)^{1/2} (d^T d)^{1/2}. \quad (48)$$

Desde que a matriz resultante do produto $\left(\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \right)$ é simétrica, de ordem p , com autovalores $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ e autovetores e_1, e_2, \dots, e_p , então

$$\max_{c \neq 0} \frac{c^T \left(\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \right) c}{c^T c} = \lambda_1,$$

que produz

$$c^T \left(\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \right) c = \lambda_1 c^T c,$$

onde λ_1 é o maior autovalor de $\left(\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \right)$, e c o autovetor normalizado e_1 correspondente, ou seja,

$$e_1^T \left(\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \right) e_1 = \lambda_1 e_1^T e_1,$$

e portanto,

$$e_1^T \left(\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \right) e_1 = \lambda_1. \quad (49)$$

Substituindo este resultado em 2.48, tem-se

$$e_1^T \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} d \leq (\lambda_1)^{1/2} (d^T d)^{1/2},$$

que por sua vez substituído em 2.47, resulta

$$\max_{a,b} \text{Corr}(U_1, V_1) = \frac{e_1^T \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} d}{\sqrt{e_1^T e_1} \cdot \sqrt{d^T d}} = \frac{(\lambda_1)^{1/2} (d^T d)^{1/2}}{1 \cdot (d^T d)^{1/2}} = \sqrt{\lambda_1},$$

ou

$$\max_{a,b} \text{Corr} (a_1^T X , b_1^T Y) = \sqrt{\lambda_1} = \rho_1^*. \quad (50)$$

Verifica-se ainda que a igualdade em 2.48 ocorre quando

$$e_1^T \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} d = (\lambda_1)^{1/2} (d^T d)^{1/2},$$

e isto assegura que d é proporcional a

$$\left(\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} e_1 \right). \quad (51)$$

A maximização da correlação indicada por 2.50 é atingida para $a = \Sigma_{11}^{-1/2}$, $c = \Sigma_{11}^{-1/2} e_1$ e b proporcional a $\left[\Sigma_{22}^{-1/2} \Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} e_1 \right]$, uma vez que $b = \Sigma_{22}^{-1/2} d$ e d é proporcional a 2.51, onde o sinal é selecionado para dar uma correlação positiva. Considerando $b = \Sigma_{22}^{-1/2} f_1$. Agora partindo expressão 2.49, tem-se

$$\left(\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \right) e_1 = \lambda_1 e_1,$$

se multiplicar ambos os membros por $\left(\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} \right)$, obtém-se

$$\begin{aligned} \left(\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} \right) \left(\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \right) e_1 &= \lambda_1 \left(\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} \right) e_1 \\ \Sigma_{22}^{-1/2} \Sigma_{21} \underbrace{\Sigma_{11}^{-1/2} \Sigma_{11}^{-1/2}} &\underbrace{\Sigma_{12} \Sigma_{22}^{-1}} \Sigma_{21} \Sigma_{11}^{-1/2} e_1 = \lambda_1 \left(\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} \right) e_1 \\ \Sigma_{22}^{-1/2} \Sigma_{21} \underbrace{\Sigma_{11}^{-1}} &\underbrace{\Sigma_{12} \Sigma_{22}^{-1/2} \Sigma_{22}^{-1/2}} \Sigma_{21} \Sigma_{11}^{-1/2} e_1 = \lambda_1 \left(\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} \right) e_1 \end{aligned}$$

$$\left[\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1/2} \right] \left(\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} e_1 \right) = \lambda_1 \left(\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} e_1 \right) \quad (52)$$

logo, se (λ_1, e_1) é um par de autovalor e autovetor de $\left(\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \right)$, então (λ_1, f_1) é um par de autovalor e autovetor de $\left[\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1/2} \right]$, onde f_1 é a forma normalizada de $\left(\Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1/2} e_1 \right)$, ao qual, d é proporcional; o sinal de f_1 é escolhido

para dar uma correlação positiva. Assim tem-se demonstrado que

$$U_1 = a_1^T X = \left(\Sigma_{11}^{-1/2} e_1 \right)^T X = e_1^T \Sigma_{11}^{-1/2} X$$

e

$$V_1 = b_1^T Y = \left(\Sigma_{22}^{-1/2} f_1 \right)^T Y = f_1^T \Sigma_{22}^{-1/2} Y.$$

U_1 e V_1 constituem o primeiro par de variáveis canônicas e sua correlação é

$$\rho_1^* = \sqrt{\lambda_1}.$$

Conclui-se ainda que

$$\begin{aligned} U_1 &= a_1^T X \\ \text{Var}(U_1) &= a_1^T \text{Var}(X) a_1 \\ \text{Var}(U_1) &= a_1^T \Sigma_{11} a_1 \\ \text{Var}(U_1) &= \left(\Sigma_{11}^{-1/2} e_1 \right)^T \Sigma_{11} \left(\Sigma_{11}^{-1/2} e_1 \right) \\ \text{Var}(U_1) &= e_1^T \Sigma_{11}^{-1/2} \Sigma_{11} \Sigma_{11}^{-1/2} e_1 \\ \text{Var}(U_1) &= e_1^T I e_1 = e_1^T e_1 \\ \text{Var}(U_1) &= 1 \end{aligned}$$

e de maneira similar

$$\text{Var}(V_1) = 1.$$

Observa-se também que U_1 e uma combinação linear arbitrária $a^T X = c^T \Sigma_{11}^{-1/2} X$ não são correlatas, se

$$0 = \text{Cov} \left(U_1, \left[c^T \Sigma_{11}^{-1/2} X \right] \right) = U_1 \Sigma_{11} \left[c^T \Sigma_{11}^{-1/2} X \right]^T = e_1^T \Sigma_{11}^{-1/2} \Sigma_{11} \Sigma_{11}^{-1/2} c = e_1^T c$$

ou seja, e_1^T e c são ortogonais. Na k - ésima etapa requer-se que $c \perp e_1, e_2, \dots, e_{k-1}$. A maximização

$$\max_{c \perp e_1, e_2, \dots, e_{k-1}} \frac{c^T B c}{c^T c} = \lambda_k \quad (c = e_k, k = 1, 2, \dots, p-1)$$

produz

$$c^T \left(\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \right) c = \lambda_k c^T c$$

para $c \perp e_1, e_2, \dots, e_{k-1}$, e em 2.47, obtém-se

$$\text{Corr}(U, V) = \text{Corr}(a^T X, b^T Y) = \frac{c^T \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} d}{\sqrt{c^T c} \sqrt{d^T d}} \leq \sqrt{\lambda_1}$$

com a igualdade para $c = e_k$ ou $a = \Sigma_{11}^{-1/2} e_k$ e $b = \Sigma_{22}^{-1/2} f_k$, como antes.

Portanto, $U_k = e_k^T \Sigma_{11}^{-1/2} X$ e $V_k = f_k^T \Sigma_{22}^{-1/2} Y$ são o k -ésimo par canônico e eles tem a correlação

$$\sqrt{\lambda_k} = \rho_k^*.$$

Embora não seja requerido, explicitamente, que V_k seja não correlacionado, tem-se

$$\text{Cov}(V_k, V_L) = f_k^T \Sigma_{22}^{-1/2} \Sigma_{22} \Sigma_{22}^{-1/2} f_L = f_k^T f_L = 0, \text{ se } k \neq L \leq p.$$

Também,

$$\text{Cov}(U_k, V_L) = e_k^T \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} f_L = 0, \text{ se } k \neq L \leq p,$$

desde que f_k^T é um múltiplo de $\left(e_k^T \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} \right)$ por 2.52.

Se as variáveis originais componentes do vetor Z forem padronizadas, obtemos o vetor Z^S . A determinação de Z^S pode ser feita efetuando-se o produto da inversa da matriz dos desvios padrão $(D^{1/2})^{-1}$ pelo vetor $(Z - \mu)$, já definidos anteriormente. Então,

$$Z^S = (D^{1/2})^{-1} (Z - \mu),$$

que produz,

$$Z^S = \begin{pmatrix} \frac{1}{\sqrt{\sigma_{11}}} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sqrt{\sigma_{pp}}} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \frac{1}{\sqrt{\sigma_{p+1p+1}}} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \frac{1}{\sqrt{\sigma_{p+qp+q}}} \end{pmatrix} \begin{pmatrix} X_1 - \mu_{x1} \\ \vdots \\ X_p - \mu_{xp} \\ Y_1 - \mu_{y1} \\ \vdots \\ Y_q - \mu_{yq} \end{pmatrix},$$

ou seja,

$$Z^S = \begin{pmatrix} \frac{X_1 - \mu_{x1}}{\sqrt{\sigma_{11}}} \\ \vdots \\ \frac{X_p - \mu_{xp}}{\sqrt{\sigma_{pp}}} \\ \frac{Y_1 - \mu_{y1}}{\sqrt{\sigma_{p+1p+1}}} \\ \vdots \\ \frac{Y_q - \mu_{yq}}{\sqrt{\sigma_{p+qp+q}}} \end{pmatrix}.$$

Irá se usar também as notações

$$Z^S = \begin{pmatrix} Z_{X_1}^S \\ \vdots \\ Z_{X_p}^S \\ Z_{Y_1}^S \\ \vdots \\ Z_{Y_q}^S \end{pmatrix} = \begin{pmatrix} Z_X^S \\ Z_Y^S \end{pmatrix}.$$

Assim, a matriz das covariâncias de Z^S será dada por

$$Cov(Z^S) = \rho = (D^{1/2})^{-1} \Sigma (D^{1/2})^{-1},$$

que produz a matriz covariância $Cov(Z^S)$, dada por :

$$\rho = \begin{pmatrix} 1 & \cdots & \rho_{1p} & \rho_{1p+1} & \cdots & \rho_{1p+q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \rho_{1p} & \cdots & 1 & \rho_{pp+1} & \cdots & \rho_{pp+q} \\ \rho_{1p+1} & \cdots & \rho_{pp+1} & 1 & \cdots & \rho_{p+1p+q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \rho_{1p+q} & \cdots & \rho_{pp+q} & \rho_{p+1p+q} & \cdots & 1 \end{pmatrix} = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix}$$

Agora, as variáveis canônicas podem ser rerepresentadas por :

$$U_k = a_k^T Z_X^S = e_k^T \rho_{11}^{-1/2} Z_X^S \quad (53)$$

e

$$V_k = b_k^T Z_Y^S = f_k^T \rho_{22}^{-1/2} Z_Y^S .$$

Aqui, $Cov(Z_X^S) = \rho_{11}$, $Cov(Z_Y^S) = \rho_{22}$, $Cov(Z_X^S, Z_Y^S) = \rho_{12} = \rho_{21}^T$, e e_k e f_k são os autovetores de $\begin{pmatrix} \rho_{11}^{-1/2} & \rho_{12} \rho_{22}^{-1} & \rho_{21} \rho_{11}^{-1/2} \end{pmatrix}$ e $\begin{pmatrix} \rho_{22}^{-1/2} & \rho_{21} \rho_{11}^{-1} & \rho_{12} \rho_{22}^{-1/2} \end{pmatrix}$, respectivamente.

As correlações canônicas, ρ_k^* , satisfazem

$$Corr(U_k, V_k) = \rho_k^*, K = 1, 2, \dots, p$$

onde $\rho_1^{*2} \leq \rho_2^{*2} \leq \dots \leq \rho_p^{*2}$ são os autovalores não nulos da matriz $\begin{pmatrix} \rho_{11}^{-1/2} & \rho_{12} \rho_{22}^{-1} & \rho_{21} \rho_{11}^{-1/2} \end{pmatrix}$, ou equivalentemente de $\begin{pmatrix} \rho_{22}^{-1/2} & \rho_{21} \rho_{11}^{-1} & \rho_{12} \rho_{22}^{-1/2} \end{pmatrix}$. Nota-se que

$$a_k^T (X - \mu_x) = a_{k1} (X_1 - \mu_{x1}) + \dots + a_{kp} (X_p - \mu_{xp}),$$

ou seja,

$$a_k^T (X - \mu_x) = a_{k1} \sqrt{\sigma_{11}} \frac{(X_1 - \mu_{x1})}{\sqrt{\sigma_{11}}} + \dots + a_{kp} \sqrt{\sigma_{pp}} \frac{(X_p - \mu_{xp})}{\sqrt{\sigma_{pp}}}$$

onde $\sigma_{ii} = Var(X_i)$, $i = 1, 2, \dots, p$. Portanto, os coeficientes canônicos para as variáveis padronizadas, $Z_{X_i}^S = \frac{(X_i - \mu_{xi})}{\sqrt{\sigma_{ii}}}$, são simplesmente relacionados com os coeficientes canônicos aderidos às variáveis originais X_i . Especificamente, se a_k^T é o vetor de coe-

ficientes para k – *ésima* variável canônica U_k , então $a_k^T D_{11}^{1/2}$ é o vetor de coeficientes para a k – *ésima* variável canônica construído a partir das variáveis padronizadas Z_X^S . Aqui, $D_{11}^{1/2}$ é uma matriz diagonal com i – *ésimo* elemento diagonal igual $\sqrt{\sigma_{ii}}$. Similarmente, $b_k^T D_{22}^{1/2}$ é o vetor de coeficientes para as variáveis canônicas construído a partir do conjunto de variáveis padronizadas Z_Y^S . Neste caso, $D_{22}^{1/2}$ é uma matriz diagonal com i – *ésimo* elemento igual a $\sqrt{\sigma_{ii}} = \sqrt{Var(Y_i)}$. As correlações canônicas são inalteradas com a padronização, embora a escolha dos vetores de coeficientes, a_k e b_k não será única se $\rho_k^{*2} = \rho_{k+1}^{*2}$.

O relacionamento entre os coeficientes canônicos das variáveis padronizadas e os coeficientes canônicos das variáveis originais seguem a especial estrutura da matriz $\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2}$ (ou $\rho_{11}^{-1/2} \rho_{12} \rho_{22}^{-1} \rho_{21} \rho_{11}^{-1/2}$). Por exemplo, se a_k^T é o vetor de coeficientes para a k – *ésima* principal componente obtida de Σ , então $a_k^T (Z - \mu) = a_k^T D^{1/2} Z^S$, mas nós não podemos inferir que $a_k^T D^{1/2}$ é o vetor de coeficientes para k – *ésima* principal componente derivada de ρ .

Capítulo 3 - Material e Método

3.1 Um Novo Método Numérico para o Cálculo de Autovalor e Autovetor de Produto de Matrizes

3.1.1 Introdução

Como já foi visto na seção anterior, a *ANÁLISE DE CORRELAÇÃO CANÔNICA* requer a determinação dos *autovalores* e *autovetores* do produto de matrizes, antes representado por

$$\Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} \quad (\text{ou equivalentemente, } \Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1/2})$$

e no caso da padronização das variáveis originais, por

$$\rho_{11}^{-1/2} \rho_{12} \rho_{22}^{-1} \rho_{21} \rho_{11}^{-1/2} \quad (\text{ou equivalentemente, } \rho_{22}^{-1/2} \rho_{21} \rho_{11}^{-1} \rho_{12} \rho_{22}^{-1/2}).$$

Em qualquer caso, o problema de correlação canônica é formulado como um problema de determinar o autosistema do produto de cinco matrizes, o qual doravante passará a ser representado por $m = m_1 m_2 m_3 m_4 m_5$, onde m é uma matriz simétrica e definida positiva. Sabe-se que se a matriz m é simétrica e positiva definida, o número de condição do problema de determinar um autovalor de m pode ser cotado superiormente, ou seja, o problema é bem-condicionado. Infelizmente este resultado não é válido para o correspondente problema de determinar o autovetor, pois os autovetores não são funções contínuas das entradas de uma matriz. De fato, pode-se mostrar que, até mesmo para uma matriz diagonal, a

inserção de pequenas perturbações em seus elementos poderá causar uma grande variação em seus autovetores. Isto ocorre freqüentemente com autovetores associados a autovalores com algum tipo de agrupamento, porém observa-se que também pode ocorrer em casos em que os autovetores estão associados à autovalores distintos e bem espaçados. Este fato pode trazer sérios problemas na interpretação dos resultados da análise da correlação canônica entre dois conjuntos de variáveis. A razão para tais problemas é que as componentes e_i e f_i (sendo cada f_i proporcional a $\Sigma_{22}^{-1/2}\Sigma_{21}\Sigma_{11}^{-1}e_i$), respectivamente, dos autovetores e_k^T e f_k da matriz m , estão estreitamente relacionadas com as componentes dos vetores de coeficientes das combinações lineares, a e b , eis que, como já visto

$$a = \Sigma_{11}^{-1/2} e_k \quad \text{e} \quad b = \Sigma_{22}^{-1/2} f_k.$$

E como, a máxima correlação desejada

$$\max_{a,b} \text{Corr}(U, V) = \rho_1^*,$$

é atingida pelo primeiro par de variáveis canônicas,

$$U_1 = a_1^T X = e_1^T \Sigma_{11}^{-1/2} X \quad \text{e} \quad V_1 = b_1^T Y = f_1^T \Sigma_{22}^{-1/2} Y,$$

então, sendo o problema de calcular os autovetores mal-condicionado, pode-se encontrar valores para os coeficientes das combinações lineares muito diferentes dos verdadeiros. Isto

é, em sistema produtivo, a interpretação dos resultados poderia levar a potencialização de ações que na realidade deveriam ser reprimidas e vice-versa e dessa forma a tomar atitudes de gerenciamento e controle, totalmente equivocadas. Sendo assim, apresenta-se a seguir novas alternativas que procuram pelo menos minimizar o problema de mal-condicionamento do cálculo dos autovetores da matriz $m = m_1 m_2 m_3 m_4 m_5$.

3.1.2 Caracterização da Matriz Produto (m)

Tendo em vista a motivação deste trabalho estar centrada na *análise da correlação canônica*, será considerado aqui o problema de determinar o autosistema da matriz $m = m_1 m_2 m_3 m_4 m_5$, onde m é uma matriz *simétrica e definida positiva*. Será considerado ainda que:

- (i) a matriz m_5 é igual *transposta* da matriz m_1 , $m_5 = m_1^T$,
- (ii) a matriz m_4 é igual *transposta* da matriz m_2 , $m_4 = m_2^T$ e
- (iii) todas as matrizes são *não singulares* e m_3 é *simétrica e definida positiva*.

3.1.3 Construção da Matriz de Cinco Blocos (M)

Seja a matriz M de cinco blocos definida por

$$M = \begin{pmatrix} & & & & m_5 \\ m_4 & & & & \\ & m_3 & & & \\ & & m_2 & & \\ & & & m_1 & \end{pmatrix},$$

onde m_1, m_2, m_3, m_4 e m_5 são matrizes como definidas anteriormente, então existe um estreito relacionamento entre os autovalores de M e os autovalores de m , como será demonstrado a seguir.

TEOREMA: Se λ é um autovalor de M , então λ^5 é um autovalor de m .

DEMONSTRAÇÃO: Sendo λ um autovalor de M , então existe um vetor

$$x^T = (x_1 \ x_2 \ x_3 \ x_4 \ x_5)$$

não nulo tal que

$$\begin{pmatrix} & & & & m_5 \\ m_4 & & & & \\ & m_3 & & & \\ & & m_2 & & \\ & & & m_1 & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}.$$

Desde que $m_5 x_5 = \lambda x_1$, $m_4 x_1 = \lambda x_2$, $m_3 x_2 = \lambda x_3$, $m_2 x_3 = \lambda x_4$, $m_1 x_4 = \lambda x_5$, tomando a equação $m_5 x_5 = \lambda x_1$ e multiplicando-a, à esquerda, consecutivamente por m_4 , m_3 , m_2 e m_1 , obtém-se

$$\begin{aligned} m_4 m_5 x_5 &= \lambda (m_4 x_1) \\ m_4 m_5 x_5 &= \lambda^2 x_2 \\ m_3 m_4 m_5 x_5 &= \lambda^2 (m_3 x_2) \\ m_3 m_4 m_5 x_5 &= \lambda^3 x_3 \\ m_2 m_3 m_4 m_5 x_5 &= \lambda^3 (m_2 x_3) \\ m_2 m_3 m_4 m_5 x_5 &= \lambda^4 x_4 \\ m_1 m_2 m_3 m_4 m_5 x_5 &= \lambda^4 (m_1 x_4), \end{aligned}$$

e finalmente

$$m_1 m_2 m_3 m_4 m_5 x_5 = \lambda^5 x_5.$$

RESULTADOS:

- A matriz M tem pelo menos um autovalor real.

- Se λ é um autovalor de M , então λ^5 é um autovalor de m . Como os autovalores de m são todos reais, então λ^5 é um número real e, portanto, λ é real.
- Como $p(\lambda) = \det(M - \lambda I)$ tem grau $5n$, e como todo autovalor de M é real, então M tem no máximo n autovalores distintos, cada um desses autovalores é um real, e satisfaz e satisfaz a relação $\lambda^5 = \alpha$, onde α é um autovalor de m .

Uma questão natural que aparece é, se dado um autovalor α qualquer de m , existe um autovalor λ de M tal que $\alpha = \lambda^5$. Se isso for verdade então, todo o autovalor de m pode ser encontrado calculando-se os autovalores de M , Em particular, pode-se obter o máximo autovalor de m por meio da matriz M . O teorema a seguir dá uma resposta a essa questão.

TEOREMA: Se α é um autovalor de m então existe um autovalor de λ de M tal que $\lambda^5 = \alpha$.

DEMONSTRAÇÃO: Seja α um autovalor de m e seja v o autovetor associado, isto é,

$$mv = \alpha v.$$

Agora, seja λ um autovalor de M associado ao autovetor $x = (x_1, x_2, x_3, x_4, x_5)$. Então o vetor $z = (z_1 \ z_2 \ z_3 \ z_4 \ z_5)^T$ definido como:

$$\begin{aligned} z_1 &= \frac{1}{\lambda} m_5 v; \\ z_2 &= \frac{1}{\lambda^2} m_4 m_5 v; \\ z_3 &= \frac{1}{\lambda^3} m_3 m_4 m_5 v; \\ z_4 &= \frac{1}{\lambda^4} m_2 m_3 m_4 m_5 v; \\ z_5 &= v, \end{aligned}$$

é um autovetor de M associado ao autovalor λ tal que $\lambda^5 = \alpha$, e portanto, dado qualquer autovalor α de M , existe um autovalor β de M , com $\beta^5 = \alpha$.

$$\begin{pmatrix} & & & & m_5 \\ m_4 & & & & \\ & m_3 & & & \\ & & m_2 & & \\ & & & m_1 & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$$

Então

$$\begin{aligned} m_5 x_5 &= \lambda x_1; \\ m_4 x_1 &= \lambda x_2; \\ m_3 x_2 &= \lambda x_3; \\ m_2 x_3 &= \lambda x_4 \text{ e} \\ m_1 x_4 &= \lambda x_5 v. \end{aligned}$$

Como $m_5 v = \lambda x_1$, então

$$\begin{aligned} x_1 &= \frac{1}{\lambda} m_5 v; \\ x_2 &= \frac{1}{\lambda^2} m_4 m_5 v; \\ x_3 &= \frac{1}{\lambda^3} m_3 m_4 m_5 v \text{ e} \\ x_4 &= \frac{1}{\lambda^4} m_2 m_3 m_4 m_5 v. \end{aligned}$$

Fazendo-se as substituições adequadas, obtém-se

$$\begin{aligned} m_4 x_1 &= \frac{1}{\lambda} m_4 m_5 v \Rightarrow m_4 m_5 v = \lambda^2 x_2; \\ m_3 x_2 &= \frac{1}{\lambda^2} m_3 m_4 m_5 v \Rightarrow m_3 m_4 m_5 v = \lambda^3 x_3; \\ m_2 x_3 &= \frac{1}{\lambda^3} m_2 m_3 m_4 m_5 v \Rightarrow m_2 m_3 m_4 m_5 v = \lambda^4 x_4 \text{ e} \\ m_1 x_4 &= \frac{1}{\lambda^4} m_1 m_2 m_3 m_4 m_5 v \Rightarrow m_1 m_2 m_3 m_4 m_5 v = \lambda^5 v. \end{aligned}$$

Logo

$$\begin{pmatrix} & & & & m_5 \\ m_4 & & & & \\ & m_3 & & & \\ & & m_2 & & \\ & & & m_1 & \end{pmatrix} \begin{pmatrix} \frac{1}{\lambda} m_5 v \\ \frac{1}{\lambda^2} m_4 m_5 v \\ \frac{1}{\lambda^3} m_3 m_4 m_5 v \\ \frac{1}{\lambda^4} m_2 m_3 m_4 m_5 v \\ v \end{pmatrix} = \lambda \begin{pmatrix} \frac{1}{\lambda} m_5 v \\ \frac{1}{\lambda^2} m_4 m_5 v \\ \frac{1}{\lambda^3} m_3 m_4 m_5 v \\ \frac{1}{\lambda^4} m_2 m_3 m_4 m_5 v \\ v \end{pmatrix}. \quad (54)$$

3.1.5 Autosistema Relacionado com a Matriz M

Muitas propriedades da matriz M são compartilhadas com as matrizes obtidas a partir de M , usando rotações de sentido horário. Assim, constrói-se a matriz M_4

$$M_4 = \begin{pmatrix} & & & m_4 \\ m_3 & & & \\ & m_2 & & \\ & & m_1 & \\ & & & m_5 \end{pmatrix}$$

onde, dando um tratamento similar ao dado para matriz M , pode-se demonstrar que

$$m_1 m_2 m_3 m_4 m_5 x_4 = \lambda^5 x_4,$$

ou seja, tem-se a mesma relação autovalor-autovetor que antes, porém agora o autovetor é x_4 ao invés de x_5 .

Sendo a matriz M_3 ,

$$M_3 = \begin{pmatrix} & & & m_3 \\ m_2 & & & \\ & m_1 & & \\ & & m_5 & \\ & & & m_4 \end{pmatrix}$$

obtém-se agora

$$m_1 m_2 m_3 m_4 m_5 x_3 = \lambda^5 x_3.$$

Considerando a matriz M_2

$$M_2 = \begin{pmatrix} & & & m_2 \\ m_1 & & & \\ & m_5 & & \\ & & m_4 & \\ & & & m_3 \end{pmatrix}$$

de novo obtém-se o mesmo tipo de relação,

$$m_1 m_2 m_3 m_4 m_5 x_2 = \lambda^5 x_2.$$

Finalmente, considerando a matriz M_1 ,

$$M_1 = \begin{pmatrix} & & & & m_1 \\ & & & & \\ & m_5 & & & \\ & & m_4 & & \\ & & & m_3 & \\ & & & & m_2 \end{pmatrix}$$

obté-m-se

$$m_1 m_2 m_3 m_4 m_5 x_1 = \lambda^5 x_1.$$

Será verificado, mais adiante, que os autovalores de todas as matrizes M_i , $i = 1, 2, 3, 4, 5$, são sempre os mesmos e que somente os autovetores mudam. Este fato irá possibilitar obter um melhor condicionamento para o problema de autovetor.

3.1.6 Decomposição de Valor Singular de M

A decomposição de valor singular de M pode ser obtida a partir da decomposição de valor singular dos blocos m_i com $i = 1, 2, 3, 4, 5$. Seja $m_i = U_i \Sigma_i V_i$, para $i = 1, 2, 3, 4, 5$, a decomposição de valor singular de cada um dos blocos m_i de M , logo pode-se escrever

$$M = \begin{pmatrix} & & & & U_5 \Sigma_5 V_5 \\ & & & & \\ U_4 \Sigma_4 V_4 & & & & \\ & U_3 \Sigma_3 V_3 & & & \\ & & U_2 \Sigma_2 V_2 & & \\ & & & U_1 \Sigma_1 V_1 & \end{pmatrix}$$

então para a matriz U definida como

$$U = \begin{pmatrix} & & & & U_5 \\ & & & U_4 & \\ & & U_3 & & \\ & U_2 & & & \\ U_1 & & & & \end{pmatrix}$$

e para a matriz V definida como

$$V = \begin{pmatrix} & & & V_4 & \\ & & & & \\ & & V_3 & & \\ V_1 & V_2 & & & \\ & & & & V_5 \end{pmatrix}$$

tem-se

$$M = U\Sigma V^T$$

onde Σ é a matriz

$$\Sigma = \begin{pmatrix} \Sigma_1 & & & & \\ & \Sigma_2 & & & \\ & & \Sigma_3 & & \\ & & & \Sigma_4 & \\ & & & & \Sigma_5 \end{pmatrix}.$$

Desde que U e V sejam matrizes unitárias e Σ é diagonal, então $M = U\Sigma V^T$ é a decomposição de valor singular de M . Assim, a decomposição de valor singular de M pode ser obtida com um custo igual à soma dos custos para obter as decomposições de valor singular de seus blocos. Agora, observa-se que

$$M^T = \begin{pmatrix} & & & m_4^T & \\ & & & & m_3^T \\ & & & & & m_2^T \\ & & & & & & m_1^T \\ m_5^T & & & & & & \end{pmatrix} \quad (56)$$

e então MM^T será

$$MM^T = \begin{pmatrix} m_5 m_5^T & & & & \\ & m_4 m_4^T & & & \\ & & m_3 m_3^T & & \\ & & & m_2 m_2^T & \\ & & & & m_1 m_1^T \end{pmatrix}.$$

Observa-se também que para cada $i = 1, 2, 3, 4$ tem-se $M_i M_i^T$ dado por

$$M_i M_i^T = \begin{pmatrix} m_4 m_4^T & & & & \\ & m_3 m_3^T & & & \\ & & m_2 m_2^T & & \\ & & & m_1 m_1^T & \\ & & & & m_5 m_5^T \end{pmatrix},$$

$$\begin{aligned}
M_3 M_3^T &= \begin{pmatrix} m_3 m_3^T & & & & \\ & m_2 m_2^T & & & \\ & & m_1 m_1^T & & \\ & & & m_5 m_5^T & \\ & & & & m_4 m_4^T \end{pmatrix}, \\
M_2 M_2^T &= \begin{pmatrix} m_2 m_2^T & & & & \\ & m_1 m_1^T & & & \\ & & m_5 m_5^T & & \\ & & & m_4 m_4^T & \\ & & & & m_3 m_3^T \end{pmatrix} \text{ e} \\
M_1 M_1^T &= \begin{pmatrix} m_1 m_1^T & & & & \\ & m_5 m_5^T & & & \\ & & m_4 m_4^T & & \\ & & & m_3 m_3^T & \\ & & & & m_2 m_2^T \end{pmatrix}.
\end{aligned}$$

Desde o espectro de $M_i M_i^T$ é dado por

$$\lambda(M_i M_i^T) = \bigcup_{k=1}^5 \lambda(m_k m_k^T), \quad i = 1, 2, \dots, 5,$$

então, todas as matrizes $M_i M_i^T$, $i = 1, 2, 3, 4, 5$ têm o mesmo número de condição.

3.1.7 Redução do Produto $m = m_1 \cdot m_2 \cdot m_3 \cdot m_4 \cdot m_5$ ao Produto $m = B m_3 B^T$.

Seja o produto $m = m_1 m_2 m_3 m_4 m_5$ como anteriormente definido. Como $m_5 = m_1^T$ e $m_4 = m_2^T$, então pode-se escrever

$$m = m_1 \cdot m_2 \cdot m_3 \cdot m_2^T \cdot m_1^T,$$

ou seja,

$$m = (m_1 \cdot m_2) \cdot m_3 \cdot (m_1 \cdot m_2)^T,$$

então,

$$m = B \cdot m_3 \cdot B^T,$$

onde,

$$B = m_1 \cdot m_2.$$

3.1.8 Construção da Matriz de Três Blocos (W).

Com um procedimento similar ao usado na construção da matriz de cinco blocos M , pode-se agora construir a matriz W , de três blocos:

$$W = \begin{pmatrix} & & B^T \\ m_3 & & \\ & B & \end{pmatrix}.$$

TEOREMA: Se λ é um autovalor de W , então λ^3 é um autovalor de m .

DEMONSTRAÇÃO: Sendo λ um autovalor de W , então existe um vetor

$$x^T = (x_1 \quad x_2 \quad x_3)$$

não nulo tal que

$$\begin{pmatrix} & & B^T \\ m_3 & & \\ & B & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Desde que $B^T x_3 = \lambda x_1$, $m_3 x_1 = \lambda x_2$, $B x_2 = \lambda x_3$, tomando a equação $B^T x_3 = \lambda x_1$ e multiplicando-a, à esquerda, consecutivamente por m_3 e B , obtém-se

$$\begin{aligned} m_3 B^T x_3 &= \lambda (m_3 x_1) \\ m_3 B^T x_3 &= \lambda^2 x_2 \\ B m_3 B^T x_3 &= \lambda^2 (B x_2) \\ (B m_3 B^T) x_3 &= \lambda^3 x_3 \end{aligned}$$

e finalmente

$$m x_3 = \lambda^3 x_3. \tag{58}$$

3.1.9 Redução do Produto $m = m_1.m_2.m_3.m_4.m_5$ ao Produto

$$m = AA^T.$$

Seja o produto $m = m_1m_2m_3m_4m_5$ como anteriormente definido, então usando o conceito de *Raiz Quadrada de uma Matriz* (seção 2.12), pode-se reescrever este produto como

$$m = m_1.m_2.m_3^{1/2}.m_3^{1/2}.m_4.m_5.$$

Mas como já foi observado anteriormente, $m_5 = m_1^T$, $m_4 = m_2^T$ e m_3 é uma *matriz simétrica e definida positiva*, logo fazendo as substituições adequadas, obtém-se

$$m = m_1.m_2.m_3^{1/2}.(m_3^{1/2})^T.(m_2)^T.(m_1)^T,$$

ou seja,

$$m = m_1.m_2.m_3^{1/2}.(m_1.m_2.m_3^{1/2})^T,$$

então,

$$m = A.A^T,$$

onde,

$$A = m_1.m_2.m_3^{1/2}. \tag{59}$$

3.1.10 Construção da Matriz de Dois Blocos (N)

Com a mesma idéia usada na construção da matriz M de cinco blocos e da matriz W de três blocos, pode-se construir uma nova matriz simétrica N de dois blocos,

$$N = \begin{pmatrix} & A^T \\ A & \end{pmatrix}$$

onde $A = m_1.m_2.m_3^{1/2}$, cujos autovalores e autovetores estão intimamente relacionados com os autovalores e autovetores da matriz $m = m_1m_2m_3m_4m_5$.

TEOREMA: Se λ é um autovalor de N , então λ^2 é um autovalor de m .

DEMONSTRAÇÃO: Sendo λ um autovalor de N , então existe um vetor

$$x^T = (x_1 \ x_2)$$

não nulo tal que

$$\begin{pmatrix} & A^T \\ A & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Desde que $A^T x_2 = \lambda x_1$, $Ax_1 = \lambda x_2$, tomando a equação $A^T x_2 = \lambda x_1$ e multiplicando-a, à esquerda, por A , obtém-se

$$AA^T x_2 = \lambda (Ax_1)$$

$$AA^T x_2 = \lambda^2 x_2$$

e finalmente

$$mx_2 = \lambda^2 x_2. \tag{60}$$

3.1.11 Decomposição de Valor Singular da Matriz A .

Como já foi visto nas seções 2.8 e 2.9, a decomposição de valor singular da matriz A é dada por

$$A = U\Sigma V^T,$$

então,

$$m = AA^T = [U \ \Sigma \ V^T] [U \ \Sigma \ V^T]^T = U\Sigma (V^T V) \Sigma^T U^T = U [\Sigma \Sigma^T] U^T$$

$$m U = U [\Sigma \Sigma^T],$$

$$m [u_1, \dots, u_n] = [u_1, \dots, u_n] \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{pmatrix}$$

e

$$m u_i = \sigma_i^2 u_i, (i = 1, 2, \dots, n),$$

onde σ_i^2 são os autovalores de m , com os respectivos autovetores u_i dados pelas colunas de U . Esta abordagem permite achar o autosistema de m , usando apenas A e sua decomposição de valor singular, $A = U\Sigma V^T$.

3.1.12 Relação entre o Autosistema de N e a Decomposição de Valor Singular de A .

Existe uma estreita relação entre o autosistema de N e a decomposição de valor singular de A , conforme será verificado pelo teorema a seguir.

TEOREMA: Seja $N = \begin{pmatrix} & A^T \\ A & \end{pmatrix}$, onde A é uma matriz quadrada com decomposição de valor singular $A = U\Sigma V^T$, onde $U = (u_1, \dots, u_n)$, $V = (v_1, \dots, v_n)$ e $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_n)$. Então os $2n$ autovalores de N são iguais a $\pm\sigma_i$ com os autovetores unitários correspondentes $\frac{1}{\sqrt{2}} \begin{pmatrix} v_i \\ \pm u_i \end{pmatrix}$.

DEMONSTRAÇÃO: Como N é real e simétrica então N tem um conjunto completo de autovalores e autovetores e todos seus autovalores são reais. Se λ_i é um autovalor de N com autovetor correspondente w_i , então

$$Nw_i = \lambda_i w_i,$$

ou seja,

$$\begin{pmatrix} & A^t \\ A & \end{pmatrix} \begin{pmatrix} w_i^{(1)} \\ w_i^{(2)} \end{pmatrix} = \lambda_i \begin{pmatrix} w_i^{(1)} \\ w_i^{(2)} \end{pmatrix},$$

que produz,

$$\begin{cases} A^T w_i^{(2)} = \lambda_i w_i^{(1)} \\ A w_i^{(1)} = \lambda_i w_i^{(2)} \end{cases} \quad (61)$$

Seja $A = U\Sigma V^T$ a decomposição de valor singular de A , então tem-se

$$Av_i = \sigma_i u_i, \quad i = 1, 2, \dots, n \text{ e}$$

$$A^T u_i = \sigma_i v_i, \quad i = 1, 2, \dots, n.$$

Logo pode-se obter uma solução para o sistema 3.61, fazendo:

$$(a) w_i^{(1)} = v_i, w_i^{(2)} = u_i e \lambda_i = \sigma_i,$$

$$(b) w_i^{(1)} = v_i, w_i^{(2)} = -u_i e \lambda_i = -\sigma_i.$$

Donde se conclui que os autovalores de N são $\pm\sigma_i$ e os autovetores correspondentes são

$$\frac{1}{\sqrt{2}} \begin{pmatrix} v_i \\ \pm u_i \end{pmatrix}, \quad i = 1, 2, \dots, n.$$

3.2 Experimentos Numéricos

3.2.1 Geração das Matrizes m_i .

Para os experimentos numéricos, optou-se por gerar randomicamente as matrizes m_1, m_2 e m_3 , atendidas as características de cada uma. Foram geradas 1200 de cada uma das matrizes m_1, m_2 e m_3 , com ordens que variaram entre 5 e 20. Inicialmente, gerou-se aleatoriamente uma matriz p , usando o comando “randn”. O comando “randn” gera matrizes com elementos aleatórios que obedecem a uma distribuição normal com as seguintes características: média 0, desvio padrão 1 e variância 1. Em seguida foram calculados os autovalores e autovetores pelo comando “eigs”, uma rotina otimizada do MATLAB, dentro de um loop condicional para garantir que todos os autovalores da matriz p sejam diferentes de zero e dessa forma sua não-singularidade. Com a matriz p , determinou-se $m_3 = pp^T$ e, assim, obteve-se m_3 simétrica e definida positiva. Gerou-se a matriz m_1 , com os mesmos comandos usados para criar p . Foi então criada a matriz m_5 , como sendo a transposta da matriz m_1 . A matriz m_2 foi gerada com a mesma rotina usada na geração de p e m_1 e em seguida foi criada a matriz m_4 , como sendo a transposta da matriz m_2 .

3.2.2 Cálculo do Autovalor e Autovetor da Matriz

$$m = m_1 m_2 m_3 m_4 m_5.$$

A matriz m foi obtida através da multiplicação das matrizes m_1, m_2, m_3, m_4 e m_5 e, então, seus autovalores e autovetores associados foram calculados através do comando “eigs”, escolhendo apenas o maior autovalor e o autovetor a ele associado.

3.2.3 Construção das Matrizes de Blocos, M , W e N .

As construções das matrizes de blocos M , W e N foram feitas de forma similar, ou seja, utilizou-se dos comandos básicos do MATLAB para a construção de uma matriz cuidando-se de inserir zeros nas entradas da matriz que não são ocupadas pelos blocos m_1, m_2, m_3, m_4 e m_5 no caso da matriz M ; B, m_3 e B^T , no caso da matriz W e A e A^T , no caso da matriz N . As matrizes B e A foram determinadas efetuando-se, respectivamente, as multiplicações $B = m_1 \cdot m_2$ e $A = m_1 \cdot m_2 \cdot m_3^{1/2}$.

3.2.4 Cálculo dos Autovalores e Autovetores das Matrizes M , W e N .

O cálculo do maior autovalor e autovetor associado de cada uma das matrizes M , W e N seguiu uma rotina similar àquela usada no caso da matriz m .

3.2.5 Cálculo do Autovalor e Autovetor da Matriz $m = AA^T$ pela decomposição singular de A .

Tendo em vista as características do produto AA^T , o maior autovalor e autovetor associado de m foram aqui calculados através da decomposição de valor singular de A . A decomposição de valor singular de A foi feita utilizando-se o comando “svd” do MATLAB. Assim, o maior autovalor de m foi determinado elevando-se ao quadrado o maior valor singular de A , e o autovetor, associado àquele maior autovalor, foi tomado como sendo a coluna da matriz U correspondente ao maior valor singular na decomposição de valor singular $A = U\Sigma V^T$.

3.2.6 Perturbação das Matrizes

Por conta do interesse em verificar o desempenho das matrizes m , M , W , N e A , na obtenção do autosistema de m , introduziram-se perturbações aleatórias que variaram da ordem de 10^{-5} à 10^{-1} , em cada uma das matrizes m , M , W , N e A . Em seguida, calcularam-se o maior autovalor e autovetor associado destas matrizes, já perturbadas, aplicando-se os mesmos métodos utilizados nas matrizes não perturbadas.

3.2.7 Precisão Numérica

O software MATLAB trabalha com precisão numérica na ordem da 16^a casa decimal para os cálculos numéricos, no entanto optou-se por apresentar os resultados com a impressão até a 6^a casa decimal, sem prejuízos à interpretação.

Os resultados impressos referem-se aos autovetores normalizados referentes ao maior autovalor, obtidos para as matrizes m , M , W , N pela fatorização de Arnoldi (comando “eigs” do MATLAB) e, de novo, para m através da decomposição de valor singular de A . Embora tenham sido “rodados” 1200 experimentos, foram apresentados somente os resultados mais representativos.

3.2.8 Análise dos Resultados Experimentais

Como exemplo apresenta-se a seguir a análise dos resultados do experimento da matriz teste 036.

$$\begin{array}{cccc}
 \begin{array}{c} m \\ \left(\begin{array}{c} 0.603443 \\ 0.391589 \\ 0.571131 \\ -0.236334 \\ -0.316970 \end{array} \right) \\ W \end{array} &
 \begin{array}{c} m_{pert} \\ \left(\begin{array}{c} 0.599307 \\ 0.392700 \\ 0.573798 \\ -0.237505 \\ -0.317750 \end{array} \right) \\ W_{pert} \end{array} &
 \begin{array}{c} M \\ \left(\begin{array}{c} 0.603443 \\ 0.391589 \\ 0.571131 \\ -0.236334 \\ -0.316970 \end{array} \right) \\ N \end{array} &
 \begin{array}{c} M_{pert} \\ \left(\begin{array}{c} 0.606465 \\ 0.392352 \\ 0.567412 \\ -0.234639 \\ -0.318196 \end{array} \right) \\ N_{pert} \end{array} \\
 \begin{array}{c} \left(\begin{array}{c} 0.603443 \\ 0.391589 \\ 0.571131 \\ -0.236334 \\ -0.316970 \end{array} \right) \\ A \end{array} &
 \begin{array}{c} \left(\begin{array}{c} 0.550788 \\ 0.367021 \\ 0.633272 \\ -0.267718 \\ -0.298699 \end{array} \right) \\ A \end{array} &
 \begin{array}{c} \left(\begin{array}{c} -0.603443 \\ -0.391589 \\ -0.571131 \\ 0.236334 \\ 0.316970 \end{array} \right) \\ A_{pert} \end{array} &
 \begin{array}{c} \left(\begin{array}{c} -0.605566 \\ -0.392725 \\ -0.568334 \\ 0.234877 \\ 0.317627 \end{array} \right) \\ A_{pert} \end{array} \\
 \begin{array}{c} \left(\begin{array}{c} -0.603443 \\ -0.391589 \\ -0.571131 \\ 0.236334 \\ 0.316970 \end{array} \right) \\ A \end{array} &
 \begin{array}{c} \left(\begin{array}{c} 0.571022 \\ 0.369291 \\ 0.560700 \\ -0.233486 \\ -0.410680 \end{array} \right) \\ A_{pert} \end{array} &
 &
 \end{array}$$

Os vetores designados por m , M , W , N e A , referem-se ao autovetor associado ao maior autovalor de m , respectivamente, determinados por: efetuando-se o produto das cinco matrizes m_i , através da matriz de cinco blocos, através da matriz de três blocos, através da matriz de dois blocos e finalmente através da decomposição de valor singular de A . Em todos experimentos esses autovetores resultaram-se iguais, com exceção de algumas componentes que tiveram seus sinais trocados por conta do sentido de aproximação das iterações do algoritmo adotado. Os vetores indicados por m_{pert} , M_{pert} , W_{pert} , N_{pert} e A_{pert} representam os autovetores, associados aos maiores autovalores, calculados após as perturbações das matrizes e obtidos pelas cinco alternativas de cálculo do autosistema de m apresenta-

das no trabalho. A análise deve ser feita comparando-se as componentes dos autovetores calculados com e sem as perturbações inseridas. Os resultados apresentados foram escolhidos dentre aqueles que mostraram a matriz m mal-condicionada. Assim, nestes casos, observa-se que as componentes do autovetor de m com menores perturbações são as determinadas pela alternativa de se usar a matriz M de cinco blocos.

Capítulo 4 - Conclusões

- É possível calcular os autovalores e autovetores da matriz m que aparece na Análise de Correlação Canônica, sem efetuar o produto indicado, $m = m_1.m_2.m_3.m_4.m_5$.
- Os autovalores e autovetores de m podem ser obtidos através do autosistema de uma matriz de 5 blocos M , definida por:

$$M = \begin{pmatrix} & & & & m_5 \\ & & & & \\ & m_4 & & & \\ & & m_3 & & \\ & & & m_2 & \\ & & & & m_1 \end{pmatrix}.$$

- Os autovalores de M e m estão relacionados pela equação $\lambda = \alpha^5$ onde α é um autovalor de M e λ é um autovalor de m e, os autovetores de m são obtidos das últimas n componentes dos autovetores de M , onde n é a ordem das matrizes m_i ($i = 1, 2, 3, 4, 5$).
- Quando a matriz m é mal-condicionada, observa-se que o cálculo de seu autosistema é melhor condicionado quando determinado através da matriz M de cinco blocos.
- O ganho em estabilidade quando se utiliza a matriz M deve-se ao fato de que evitando o cálculo do produto $m_1.m_2.m_3.m_4.m_5$ a propagação de erros é diminuída.
- O custo de se usar a matriz M é o aumento de complexidade do cálculo nessa alternativa em relação ao cálculo direto do autosistema da matriz m efetuando-se a multiplicação $m_1.m_2.m_3.m_4.m_5$.

- Se a matriz m for estável ou medianamente estável, pode-se calcular o autosistema da matriz m através da matriz W de três blocos, N de dois blocos ou pela decomposição de valor singular da matriz A . Sendo $W = \begin{pmatrix} m_3 & B^T \\ & B \end{pmatrix}$, com $B = m_1.m_2$ e $N = \begin{pmatrix} & A^T \\ A & \end{pmatrix}$, com $A = m_1.m_2.m_3^{1/2}$.
- Os autovalores de W e m estão relacionados pela equação $\lambda = \beta^3$ onde β é um autovalor de W e λ é um autovalor de m , e os autovetores da matriz m são dados pelas últimas n componentes dos autovetores de W , sendo n a ordem das matrizes m_i ($i = 1, 2, 3, 4, 5$).
- Os autovalores de N e m estão relacionados por $\lambda = \gamma^2$ onde γ é um autovalor de N e λ é um autovalor de m , e os autovetores de m são dados pelas últimas n componentes dos autovetores de N .
- Sendo $A = U\Sigma V^T$ a decomposição de valor singular de A e σ_i um valor singular de A então, σ_i^2 é um autovalor de m e, os autovetores da matriz m são dados pelas colunas u_i da matriz U .

- As perturbações decorrentes da má qualidade dos dados de entrada são mais significativas que as perturbações decorrentes da instabilidade dos algoritmos utilizados.

4.1 Propostas para Trabalhos Futuros

- É importante, na continuidade dos estudos, que se pesquise a propagação de erros, tanto na construção como no cálculo do autossistema das matrizes m , M , W e N . Além disso, estudar qual o impacto de instabilidade nas matrizes causado pelas perturbações inseridas por imprecisão dos dados ou por instabilidade do algoritmo adotado.
- Outra sugestão é o estudo da aplicação da análise de correlção canônica para identificar e quantificar as associações lineares entre os conjuntos de variáveis aleatórias envolvidas no processo de usinagem de Misturas Asfálticas para Pavimentação.

Referências Bibliográficas

- [STR93]STRANG,GILBERT. *Linear Algebra and Its Applications*. Second Edition. London. Academic Press, Inc.,1976.
- [GOL96]GOLUB, GENE H.; LOAN CHARLES F. VAN. *Matrix Computations*. Third Edition. London. The Johns Hopkins University Press Ltd.,1996.
- [NOB86]NOBLE,BEN. *Applied Linear Algebra*. Second Edition. New Jersey. Prentice Hall, Inc., 1986
- [LAY97]LAY, DAVID C..*Linear Algebra and Its Applications*. Maryland. Addison Wesley Longman, Inc., 1997. .
- [JON88]JONSON, RICHARD A.;WICHERN, DEAN W. *Applied Multivariate Statistic Analisis*. Second Edition. New Jersey. Prentice Hall International, 1988.
- [MON96]MONTGOMERY DOUGLAS C.. *Design and Analysis of Experiments*. Fourth Edition. New York. John Wiley & Sons, 1996.
- [FEY99]FEY NETO, EMÍLIO R.*Análise de Correlação Canônica Aplicada em Sistemas de Produção Contínuo*, Dissertação de Mestrado, Curitiba, PPGIA, PUCPR, 1999.
- [LEL02]LELUDAK, JORGE ASSAD.*Um Novo Método para Cálculo dos Estimadores dos Parâmetros da Correlação Canônica*, Dissertação de Mestrado, Curitiba, PPGIA, PUCPR, 2002.
- [SAM01]SAMPAIO, RAIMUNDO J. B. de; CANDIDO, MARCO A. B.*An Stable Algorithm to Calculate Canonical Correlation*, Article, Curitiba, PPGIA, PUCPR, 2001.
- [FEY00]FEY NETO, EMÍLIO; CHAVES NETO, ANSELMO; SAMPAIO, RAIMUNDO J. B. de; ZUGE, MARLENE. *Análise de Correlação Canônica Aplicada a Sistemas Contínuos de Produção*, Article In X CLAIO, México, 2000.
- [SAM00]SAMPAIO, RAIMUNDO J. B. de; FEY NETO, EMÍLIO; CHAVES NETO, ANSELMO. *Canonical Correlation Analysis Applied to the Verification of Important Variables of the Stages of a Continuos Process Associate to th Quality of Final Product*. IV INDUSTRY APPLICATIONS CONFERENCE, Porto Alegre, 2000.
- [HAR01]HARTFIEL, DARALD J.. *Matrix Theory and Applications with MATLAB*. London. CRC Press LLC, 2001.

[PEN00]PENNY, JOHN ; LINDFIELD, GEORGE. *Numerical Methods Using MATLAB*. New Jersey. Prentice Hall, Inc., 2000.

[KNI00]KNIGHT, ANDREW. *Basics of MATLAB and Beyond*. London. CRC Press LLC, 2000.

Apêndice A - Rotina Para os Experimentos Numéricos

A seguir é apresentada a rotina, comentada, utilizada no MATLAB para produzir os experimentos numéricos.

```
clear;

for mat=1:2

% Define a dimensão da matriz quadrada mk (por exemplo de 5a ordem)

Lm=5;

% Define o nome do arquivo de saída

FILENAME1=sprintf('%s%03d%s%02d%s', 'matriz', mat, '-', Lm, 'avn.txt');

fid1=fopen(FILENAME1, 'w');

% Define a perturbação da matriz quadrada mk (por exemplo da ordem de 10-5)

pert=13e-5;

% Inicializa a variável de teste para verificar se a matriz

% randômica criada atende às condições para a Matriz p

testep=1;

% Cria um loop Condicional para que a matriz p seja não-singular

% e desta forma m3=p*TRANSPOSTA(p) seja positiva definida

while testep~=0

% Define uma Matriz quadrada com valores randômicos para

% criar a matriz p
```

```
p=randn(Lm);

% Define a opção para não plotar as saídas intermediárias do
% cálculo dos autovalores e autovetores

opts.disp=0;

% Calcula autovetores V3 e autovalores D3 para a Matriz m3

[Vp,Dp]=eigs(p,Lm,'LR',opts);

% Teste os autovalores de m3 para garantir que esta seja
% uma Matriz positiva definida

ttp=0;

for i=1:Lm

if Dp(i,i)~=0.0;

ttp=ttp+1;

end

end

if ttp==Lm;

%disp('OK - p Não Singular');

testep=0;

end

% Finaliza o condicional quando a Matriz p é não-singular

end

% Gear m3 positiva definida

m3=p*p';% Inicializa a variável de teste para verificar se a matriz
```

```
%    randômica criada atende às condições para a Matriz m1

testem1=1;

%    Cria um loop Condicional para que a matriz m1 seja

% não-singular

while testem1~=0

% Define uma Matriz quadrada com valores randômicos para

% criar a matriz m1

m1=randn(Lm);

% Define a opção para não plotar as saídas intermediárias do

%    cálculo dos autovalores e autovetores

opts.disp=0;

% Calcula autovetores V1 e autovalores D1 para a Matriz m1

[V1,D1]=eigs(m1,Lm,'LR',opts);

% Teste os autovalores de m1 para garantir que esta seja

%    uma Matriz Não Singular

ttm1=0;

for i=1:Lm

if D1(i,i)~=0.0;

ttm1=ttm1+1;

end

end

if ttm1==Lm;
```

```
%disp('OK - m1 Não Singular');

testem1=0;

end

% Finaliza o condicional quando a Matriz m1 é não-singular.

end

% Define a matriz m5 como a transposta de m1

m5=m1';

% Inicializa a variável de teste para verificar se a matriz

%   randômica criada atende às condições para a Matriz m2

testem2=1;

%   Cria um loop Condicional para que a matriz m2 seja

%   não-singular

while testem2~=0

% Define uma Matriz quadrada com valores randômicos para

% criar a matriz m2

m2=rand(Lm);

% Define a opção para não plotar as saídas intermediárias do

%   cálculo dos autovalores e autovetores

opts.disp=0;

% Calcula autovetores V2 e autovalores D2 para a Matriz m2

[V2,D2]=eigs(m2,Lm,'LR',opts);

% Teste os autovalores de m2 para garantir que esta seja
```

```
%    uma Matriz Não Singular

ttm2=0;

for i=1:Lm

if D2(i,i)~=0.0;

ttm2=ttm2+1;

end

end

if ttm2==Lm;

%disp('OK - m2 Não Singular');

testem2=0;

end

% Finaliza o condicional quando a Matriz m2 é não-singular

end

% Define a matriz m4 como a transposta de m2

m4=m2';

% Imprime os valores das matrizes m1, m2 e m3 para verificação

fprintf(fid1,'%s\n','Matriz m1 utilizada');

for i=1:Lm

for j=1:Lm

fprintf(fid1,'%14.8f',m1(i:i,j));

end

fprintf(fid1,'\n');
```

```
end

fprintf(fid1,'%s\n','Matriz m2 utilizada');

for i=1:Lm

for j=1:Lm

fprintf(fid1,'%14.8f',m2(i:i,j:j));

end

fprintf(fid1,'\n');

end

fprintf(fid1,'%s\n','Matriz m3 utilizada');

for i=1:Lm

for j=1:Lm

fprintf(fid1,'%14.8f',m3(i:i,j:j));

end

fprintf(fid1,'\n');

end

% Calcula a Matriz m

m=m1*m2*m3*m4*m5;

% Imprime os valores da matriz m para verificação

fprintf(fid1,'%s\n','Matriz m utilizada');

for i=1:Lm

for j=1:Lm

fprintf(fid1,'%14.8f',m(i:i,j:j));
```



```

end

fprintf(fid1,'\n');

end

% Define a opção para não plotar as saídas intermediárias do

% cálculo dos autovalores e autovetores

opts.disp=0;

% Calcula o maior autovalor e autovetor associado para m

tic

[Vm,Dm]=eigs(m,1,'LR',opts);

toc

Vmmax=Vm(1:Lm,1:1);

% Imprime os autovetores e autovalores da matriz m

fprintf(fid1,'%s\n','Autovalores para a Matriz m utilizada');

fprintf(fid1,'%14.8f\n',Dm);

fprintf(fid1,'%s\n','Autovetores para a Matriz m utilizada');

fprintf(fid1,'%14.8f\n',Vm);

% Perturbando a Matriz m no elemento m(Lm:Lm,1:1) com uma perturbação

% pert=13e-5.

mpert=m;

mpert(1:1,Lm:Lm)=mpert(1:1,Lm:Lm)+pert;

% Calcula o maior autovalor e autovetor associado para mpert

[Vmpert,Dmpert]=eigs(mpert,1,'LR',opts);

```

```

Vmpertmax=Vmpert(1:Lm,1:1);

fprintf(fid1,'%s\n','Autovalores para a Matriz m perturbada utilizada');

fprintf(fid1,'%14.8f\n',Dmpert);

fprintf(fid1,'%s\n','Autovetores para a Matriz m perturbada utilizada');

fprintf(fid1,'%14.8f\n',Vmpert);

% Cria uma Matriz de zeros, m0

m0=zeros(Lm);

% Define a dimensão da Matriz em bloco M

LM = 5*Lm;

% Cria a Matriz em bloco M

M=[m0 m0 m0 m0 m5;m4 m0 m0 m0 m0;m0 m3 m0 m0 m0;m0 m0 m2 m0 m0;m0 m0 m0 m1 m0];

% Define a opção para não plotar as saídas intermediárias do

% cálculo dos autovalores e autovetores

opts.disp=0;

% Calcula o maior autovalor e autovetor associado para M

tic

[VM,DM]=eigs(M,1,'LR',opts);

toc

VMmax=VM((4*Lm+1):LM,1:1);

fprintf(fid1,'%s\n','Autovalores para a Matriz M utilizada');

fprintf(fid1,'%14.8f\n',DM);

fprintf(fid1,'%s\n','Autovetores para a Matriz M utilizada');

```

```

fprintf(fid1,'%14.8f\n',VM);

% Perturbando a Matriz M no elemento m(LM:LM,1:1) com perturbação

% pert=13e-5

Mpert=M;

Mpert(1:1,LM:LM)=Mpert(1:1,LM:LM)+pert;

% Calcula o maior autovalor e autovetor associado para Mpert

[VMpert,DMpert]=eigs(Mpert,1,'LR',opts);

VMpertmax=VMpert((4*Lm+1):LM,1:1);

fprintf(fid1,'%s\n','Autovalores para a Matriz Mpert utilizada');

fprintf(fid1,'%14.8f\n',DMpert);

fprintf(fid1,'%s\n','Autovetores para a Matriz Mpert utilizada');

fprintf(fid1,'%14.8f\n',VMpert);

LB=3*Lm;

B=m1*m2;

% OBS.: m=B*m3*TRANSPOSTA(B)

Bt=B';

W=[m0 m0 Bt;m3 m0 m0;m0 B m0];

% Calcula o maior autovalor e autovetor associado para W

tic

[VW,DW]=eigs(W,1,'LR',opts);

toc

VWmax=VW(2*Lm+1:LB,1:1);

```

```

fprintf(fid1,'%s\n','Autovalores para a Matriz W utilizada');

fprintf(fid1,'%14.8f\n',DW);

fprintf(fid1,'%s\n','Autovetores para a Matriz W utilizada');

fprintf(fid1,'%14.8f\n',VW);

% Perturbando a Matriz W no elemento W(LB:LB,1:1) com perturbação

% pert=13e-5

Wpert=W;

Wpert(1:1, LB:LB)=Wpert(1:1, LB:LB)+pert;

% Calcula o maior autovalor e autovetor associado para Wpert

[VWpert,DWpert]=eigs(Wpert,1,'LR',opts);

VWpertmax=VWpert(2*Lm+1:LB,1:1);

fprintf(fid1,'%s\n','Autovalores para a Matriz Wpert utilizada');

fprintf(fid1,'%14.8f\n',DWpert);

fprintf(fid1,'%s\n','Autovetores para a Matriz Wpert utilizada');

fprintf(fid1,'%14.8f\n',VWpert);

% Extrai a raiz quadrada da matriz m3

m3r=sqrtm(m3);

A=m1*m2*m3r;

LA=2*Lm;

At=A';

N=[m0 At; A m0];

% Calcula o maior autovalor e autovetor associado para N

```

```

tic

[VN,DN]=eigs(N,1,'LA',opts);

toc

VNmax=VN(Lm+1:LA,1:1);

fprintf(fid1,'%s\n','Autovalores para a Matriz N utilizada');

fprintf(fid1,'%14.8f\n',DN);

fprintf(fid1,'%s\n','Autovetores para a Matriz N utilizada');

fprintf(fid1,'%14.8f\n',VN);

% Perturbando a Matriz N no elemento N(LA:LA,1:1) com perturbação

% pert=13e-5

Npert=N;

Npert(1:1,LA:LA)=Npert(1:1,LA:LA)+pert;

% Calcula o maior autovalor e autovetor associado para Npert

[VNpert,DNpert]=eigs(Npert,1,'LR',opts);

VNpertmax=VNpert(Lm+1:LA,1:1);

fprintf(fid1,'%s\n','Autovalores para a Matriz Npert utilizada');

fprintf(fid1,'%14.8f\n',DNpert);

fprintf(fid1,'%s\n','Autovetores para a Matriz Npert utilizada');

fprintf(fid1,'%14.8f\n',VNpert);

%   OBS.: m=A*TRANSPOSTA(A)

% Calcula a decomposição de valor singular para a matriz A,

%   UA matriz de autovetores ortonormais para A

```

```

%    VA matriz de autovetores ortonormais para A

%    SA matriz diagonal com valores singulares de A, sigma
tic

[UA,SA,VA]=svds(A,1);

toc

VAmatrix=UA(1:Lm,1:1);

DA=SA*SA;

fprintf(fid1,'%s\n','Autovalores para a Matriz A utilizada');

fprintf(fid1,'%14.8f\n',DA);

fprintf(fid1,'%s\n','Autovetores para a Matriz A utilizada');

fprintf(fid1,'%14.8f\n',VA);

% Perturbando a Matriz A no elemento A(Lm:Lm,1:1) com perturbação

% pert=13e-5

Apert=A;

Apert(1:1,Lm:Lm)=Apert(1:1,Lm:Lm)+pert;

% Calcula a decomposição de valor singular para a matriz Apert,

%    UApert matriz de autovetores ortonormais para Apert

%    VApert matriz de autovetores ortonormais para Apert

%    SApert matriz diagonal com valores singulares de Apert, sigma

[UApert,SApert,VApert]=svds(Apert,1);

VApertmax=UApert(1:Lm,1:1);

% Determina-se o maior autovalor de m elevando-se ao quadrado o maior valor

```

```

% singular de A

DApert=SApert*SApert;

fprintf(fid1,'%s\n','Autovalores para a Matriz Apert utilizada');

fprintf(fid1,'%14.8f\n',DApert);

fprintf(fid1,'%s\n','Autovetores para a Matriz Apert utilizada');

fprintf(fid1,'%14.8f\n',VApert);

% Normalizando os autovetores (autovetores unitários)

Vmmax=Vmmax*(1/norm(Vmmax));

fprintf(fid1,'%s\n','Autovetores normados para a Matriz m utilizada');

fprintf(fid1,'%14.8f\n',Vmmax);

Vmpertmax=Vmpertmax*(1/norm(Vmpertmax));

fprintf(fid1,'%s\n','Autovetores normados para a Matriz mpert utilizada');

fprintf(fid1,'%14.8f\n',Vmpertmax);

VMmax=VMmax*(1/norm(VMmax));

fprintf(fid1,'%s\n','Autovetores normados para a Matriz M utilizada');

fprintf(fid1,'%14.8f\n',VMmax);

VMpertmax=VMpertmax*(1/norm(VMpertmax));

fprintf(fid1,'%s\n','Autovetores normados para a Matriz Mpert utilizada');

fprintf(fid1,'%14.8f\n',VMpertmax);

VWmax=VWmax*(1/norm(VWmax));

fprintf(fid1,'%s\n','Autovetores normados para a Matriz W utilizada');

fprintf(fid1,'%14.8f\n',VWmax);

```

```
VWpertmax=VWpertmax*(1/norm(VWpertmax));  
  
fprintf(fid1,'%s\n','Autovetores normados para a Matriz W utilizada');  
  
fprintf(fid1,'%14.8f\n',VWmax);  
  
VNmax=VNmax*(1/norm(VNmax));  
  
fprintf(fid1,'%s\n','Autovetores normados para a Matriz N utilizada');  
  
fprintf(fid1,'%14.8f\n',VNmax);  
  
VNpertmax=VNpertmax*(1/norm(VNpertmax));  
  
fprintf(fid1,'%s\n','Autovetores normados para a Matriz Npert utilizada');  
  
fprintf(fid1,'%14.8f\n',VNpertmax);  
  
VAmx=VAmx*(1/norm(VAmx));  
  
fprintf(fid1,'%s\n','Autovetores normados para a Matriz A utilizada');  
  
fprintf(fid1,'%14.8f\n',VAmx);  
  
VApertmax=VApertmax*(1/norm(VApertmax));  
  
fprintf(fid1,'%s\n','Autovetores normados para a Matriz Apert utilizada');  
  
fprintf(fid1,'%14.8f\n',VApertmax);  
  
% Finaliza a impressão das saídas no arquivo  
  
fclose(fid1);  
  
% Finaliza looping para cálculo das matrizes  
clear;
```


Matriz de teste 036

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} 0.603443 \\ 0.391589 \\ 0.571131 \\ -0.236334 \\ -0.316970 \end{array} \right) & \left(\begin{array}{c} 0.599307 \\ 0.392700 \\ 0.573798 \\ -0.237505 \\ -0.317750 \end{array} \right) & \left(\begin{array}{c} 0.603443 \\ 0.391589 \\ 0.571131 \\ -0.236334 \\ -0.316970 \end{array} \right) & \left(\begin{array}{c} 0.606465 \\ 0.392352 \\ 0.567412 \\ -0.234639 \\ -0.318196 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} 0.603443 \\ 0.391589 \\ 0.571131 \\ -0.236334 \\ -0.316970 \end{array} \right) & \left(\begin{array}{c} 0.550788 \\ 0.367021 \\ 0.633272 \\ -0.267718 \\ -0.298699 \end{array} \right) & \left(\begin{array}{c} -0.603443 \\ -0.391589 \\ -0.571131 \\ 0.236334 \\ 0.316970 \end{array} \right) & \left(\begin{array}{c} -0.605566 \\ -0.392725 \\ -0.568334 \\ 0.234877 \\ 0.317627 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} -0.603443 \\ -0.391589 \\ -0.571131 \\ 0.236334 \\ 0.316970 \end{array} \right) & \left(\begin{array}{c} 0.571022 \\ 0.369291 \\ 0.560700 \\ -0.233486 \\ -0.410680 \end{array} \right) &
 \end{array}$$

Matriz de teste 065

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} -0.214448 \\ -0.672946 \\ 0.515092 \\ 0.025444 \\ 0.484963 \end{array} \right) & \left(\begin{array}{c} -0.213479 \\ -0.673071 \\ 0.515194 \\ 0.025442 \\ 0.485108 \end{array} \right) & \left(\begin{array}{c} 0.214448 \\ 0.672946 \\ -0.515092 \\ -0.025444 \\ -0.484963 \end{array} \right) & \left(\begin{array}{c} -0.214059 \\ -0.672839 \\ 0.515066 \\ 0.025379 \\ 0.485314 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} 0.214448 \\ 0.672946 \\ -0.515092 \\ -0.025444 \\ -0.484963 \end{array} \right) & \left(\begin{array}{c} 0.205563 \\ 0.669902 \\ -0.514419 \\ -0.024310 \\ -0.493718 \end{array} \right) & \left(\begin{array}{c} 0.214448 \\ 0.672946 \\ -0.515092 \\ -0.025444 \\ -0.484963 \end{array} \right) & \left(\begin{array}{c} -0.215963 \\ -0.673401 \\ 0.515203 \\ 0.025729 \\ 0.483524 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} 0.214448 \\ 0.672946 \\ -0.515092 \\ -0.025444 \\ -0.484963 \end{array} \right) & \left(\begin{array}{c} 0.219096 \\ 0.680298 \\ -0.520334 \\ -0.026118 \\ -0.466650 \end{array} \right) &
 \end{array}$$

Matriz de teste 067

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} -0.330782 \\ 0.127838 \\ 0.101145 \\ 0.856578 \\ -0.360950 \end{array} \right) & \left(\begin{array}{c} 0.331476 \\ -0.127806 \\ -0.101122 \\ -0.856356 \\ 0.360856 \end{array} \right) & \left(\begin{array}{c} 0.330782 \\ -0.127838 \\ -0.101145 \\ -0.856578 \\ 0.360950 \end{array} \right) & \left(\begin{array}{c} 0.330774 \\ -0.127788 \\ -0.101063 \\ -0.856573 \\ 0.361008 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} 0.330782 \\ -0.127838 \\ -0.101145 \\ -0.856578 \\ 0.360950 \end{array} \right) & \left(\begin{array}{c} 0.331136 \\ -0.128857 \\ -0.101657 \\ -0.856833 \\ 0.359510 \end{array} \right) & \left(\begin{array}{c} 0.330782 \\ -0.127838 \\ -0.101145 \\ -0.856578 \\ 0.360950 \end{array} \right) & \left(\begin{array}{c} -0.330310 \\ 0.126995 \\ 0.101620 \\ 0.856290 \\ -0.362226 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} -0.330782 \\ 0.127838 \\ 0.101145 \\ 0.856578 \\ -0.360950 \end{array} \right) & \left(\begin{array}{c} -0.332295 \\ 0.129057 \\ 0.101111 \\ 0.859695 \\ -0.351604 \end{array} \right) &
 \end{array}$$

Matriz de teste 098

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} 0.013336 \\ 0.016957 \\ -0.952033 \\ -0.200328 \\ -0.230298 \end{array} \right) & \left(\begin{array}{c} -0.000010 \\ 0.017242 \\ -0.952543 \\ -0.200356 \\ -0.228520 \end{array} \right) & \left(\begin{array}{c} -0.013336 \\ -0.016957 \\ 0.952033 \\ 0.200328 \\ 0.230298 \end{array} \right) & \left(\begin{array}{c} -0.014235 \\ -0.015346 \\ 0.950129 \\ 0.198326 \\ 0.239759 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} -0.013336 \\ -0.016957 \\ 0.952033 \\ 0.200328 \\ 0.230298 \end{array} \right) & \left(\begin{array}{c} 0.016143 \\ 0.015621 \\ -0.949479 \\ -0.198019 \\ -0.242432 \end{array} \right) & \left(\begin{array}{c} -0.013336 \\ -0.016957 \\ 0.952033 \\ 0.200328 \\ 0.230298 \end{array} \right) & \left(\begin{array}{c} -0.013133 \\ -0.017110 \\ 0.952269 \\ 0.200567 \\ 0.229109 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} 0.013336 \\ 0.016957 \\ -0.952033 \\ -0.200328 \\ -0.230298 \end{array} \right) & \left(\begin{array}{c} 0.005941 \\ 0.024022 \\ -0.969727 \\ -0.213113 \\ -0.116623 \end{array} \right) &
 \end{array}$$

Matriz de teste 128

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} -0.130451 \\ -0.046707 \\ 0.338593 \\ 0.545552 \\ 0.754008 \end{array} \right) & \left(\begin{array}{c} -0.129962 \\ -0.046711 \\ 0.338621 \\ 0.545585 \\ 0.754056 \end{array} \right) & \left(\begin{array}{c} -0.130451 \\ -0.046707 \\ 0.338593 \\ 0.545552 \\ 0.754008 \end{array} \right) & \left(\begin{array}{c} -0.130535 \\ -0.046690 \\ 0.338480 \\ 0.545592 \\ 0.754017 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} 0.130451 \\ 0.046707 \\ -0.338593 \\ -0.545552 \\ -0.754008 \end{array} \right) & \left(\begin{array}{c} 0.130711 \\ 0.046640 \\ -0.338275 \\ -0.545662 \\ -0.754031 \end{array} \right) & \left(\begin{array}{c} -0.130451 \\ -0.046707 \\ 0.338593 \\ 0.545552 \\ 0.754008 \end{array} \right) & \left(\begin{array}{c} -0.129903 \\ -0.046882 \\ 0.338958 \\ 0.545311 \\ 0.754103 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} 0.130451 \\ 0.046707 \\ -0.338593 \\ -0.545552 \\ -0.754008 \end{array} \right) & \left(\begin{array}{c} 0.130452 \\ 0.046707 \\ -0.338593 \\ -0.545553 \\ -0.754007 \end{array} \right) &
 \end{array}$$

Matriz de teste 134

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} -0.089750 \\ -0.060340 \\ 0.347183 \\ -0.214015 \\ -0.906623 \end{array} \right) & \left(\begin{array}{c} -0.094300 \\ -0.060272 \\ 0.347164 \\ -0.214331 \\ -0.906098 \end{array} \right) & \left(\begin{array}{c} -0.089750 \\ -0.060340 \\ 0.347183 \\ -0.214015 \\ -0.906623 \end{array} \right) & \left(\begin{array}{c} -0.089572 \\ -0.060344 \\ 0.347177 \\ -0.213940 \\ -0.906661 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} -0.089750 \\ -0.060340 \\ 0.347183 \\ -0.214015 \\ -0.906623 \end{array} \right) & \left(\begin{array}{c} -0.074033 \\ -0.061237 \\ 0.343455 \\ -0.207896 \\ -0.910817 \end{array} \right) & \left(\begin{array}{c} 0.089750 \\ 0.060340 \\ -0.347183 \\ 0.214015 \\ 0.906623 \end{array} \right) & \left(\begin{array}{c} 0.082636 \\ 0.061201 \\ -0.345551 \\ 0.212122 \\ 0.908308 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} -0.089750 \\ -0.060340 \\ 0.347183 \\ -0.214015 \\ -0.906623 \end{array} \right) & \left(\begin{array}{c} -0.087792 \\ -0.060376 \\ 0.345824 \\ -0.212961 \\ -0.907580 \end{array} \right) &
 \end{array}$$

Matriz de teste 140

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} -0.362339 \\ -0.088974 \\ 0.432902 \\ 0.492229 \\ 0.656583 \end{array} \right) & \left(\begin{array}{c} 0.359888 \\ 0.089008 \\ -0.433373 \\ -0.492765 \\ -0.657213 \end{array} \right) & \left(\begin{array}{c} 0.362339 \\ 0.088974 \\ -0.432902 \\ -0.492229 \\ -0.656583 \end{array} \right) & \left(\begin{array}{c} 0.364589 \\ 0.089163 \\ -0.431581 \\ -0.490196 \\ -0.657701 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} -0.362339 \\ -0.088974 \\ 0.432902 \\ 0.492229 \\ 0.656583 \end{array} \right) & \left(\begin{array}{c} 0.361028 \\ 0.091733 \\ -0.436605 \\ -0.492030 \\ -0.654619 \end{array} \right) & \left(\begin{array}{c} -0.362339 \\ -0.088974 \\ 0.432902 \\ 0.492229 \\ 0.656583 \end{array} \right) & \left(\begin{array}{c} 0.361507 \\ 0.087216 \\ -0.432600 \\ -0.493281 \\ -0.656686 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} 0.362339 \\ 0.088974 \\ -0.432902 \\ -0.492229 \\ -0.656583 \end{array} \right) & \left(\begin{array}{c} -0.365272 \\ -0.087963 \\ 0.437665 \\ 0.498863 \\ 0.646857 \end{array} \right) &
 \end{array}$$

Matriz de teste 168

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} 0.603262 \\ 0.205977 \\ 0.157312 \\ 0.171146 \\ -0.734582 \end{array} \right) & \left(\begin{array}{c} 0.602824 \\ 0.206056 \\ 0.157383 \\ 0.171220 \\ -0.734886 \end{array} \right) & \left(\begin{array}{c} -0.603262 \\ -0.205977 \\ -0.157312 \\ -0.171146 \\ 0.734582 \end{array} \right) & \left(\begin{array}{c} -0.603265 \\ -0.205984 \\ -0.157305 \\ -0.171140 \\ 0.734580 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} 0.603262 \\ 0.205977 \\ 0.157312 \\ 0.171146 \\ -0.734582 \end{array} \right) & \left(\begin{array}{c} -0.601910 \\ -0.203222 \\ -0.160487 \\ -0.173721 \\ 0.735166 \end{array} \right) & \left(\begin{array}{c} -0.603262 \\ -0.205977 \\ -0.157312 \\ -0.171146 \\ 0.734582 \end{array} \right) & \left(\begin{array}{c} 0.603296 \\ 0.206048 \\ 0.157250 \\ 0.171079 \\ -0.734563 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} 0.603262 \\ 0.205977 \\ 0.157312 \\ 0.171146 \\ -0.734582 \end{array} \right) & \left(\begin{array}{c} -0.603236 \\ -0.205412 \\ -0.157963 \\ -0.171865 \\ 0.734454 \end{array} \right) &
 \end{array}$$

Matriz de teste 187

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} 0.025041 \\ 0.454619 \\ -0.424459 \\ 0.560047 \\ -0.546696 \end{array} \right) & \left(\begin{array}{c} 0.024957 \\ 0.454620 \\ -0.424460 \\ 0.560048 \\ -0.546698 \end{array} \right) & \left(\begin{array}{c} 0.025041 \\ 0.454619 \\ -0.424459 \\ 0.560047 \\ -0.546696 \end{array} \right) & \left(\begin{array}{c} -0.025116 \\ -0.454686 \\ 0.424437 \\ -0.560051 \\ 0.546649 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} 0.025041 \\ 0.454619 \\ -0.424459 \\ 0.560047 \\ -0.546696 \end{array} \right) & \left(\begin{array}{c} 0.024772 \\ 0.454563 \\ -0.424488 \\ 0.559971 \\ -0.546810 \end{array} \right) & \left(\begin{array}{c} -0.025041 \\ -0.454619 \\ 0.424459 \\ -0.560047 \\ 0.546696 \end{array} \right) & \left(\begin{array}{c} -0.023696 \\ -0.454481 \\ 0.424659 \\ -0.559577 \\ 0.547196 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} 0.025041 \\ 0.454619 \\ -0.424459 \\ 0.560047 \\ -0.546696 \end{array} \right) & \left(\begin{array}{c} -0.025578 \\ -0.455036 \\ 0.424714 \\ -0.560656 \\ 0.545501 \end{array} \right) &
 \end{array}$$

Matriz de teste 228

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} -0.130451 \\ -0.046707 \\ 0.338593 \\ 0.545552 \\ 0.754008 \end{array} \right) & \left(\begin{array}{c} -0.129962 \\ -0.046711 \\ 0.338621 \\ 0.545585 \\ 0.754056 \end{array} \right) & \left(\begin{array}{c} -0.130451 \\ -0.046707 \\ 0.338593 \\ 0.545552 \\ 0.754008 \end{array} \right) & \left(\begin{array}{c} -0.130535 \\ -0.046690 \\ 0.338480 \\ 0.545592 \\ 0.754017 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} 0.130451 \\ 0.046707 \\ -0.338593 \\ -0.545552 \\ -0.754008 \end{array} \right) & \left(\begin{array}{c} 0.130711 \\ 0.046640 \\ -0.338275 \\ -0.545662 \\ -0.754031 \end{array} \right) & \left(\begin{array}{c} -0.130451 \\ -0.046707 \\ 0.338593 \\ 0.545552 \\ 0.754008 \end{array} \right) & \left(\begin{array}{c} -0.129903 \\ -0.046882 \\ 0.338958 \\ 0.545311 \\ 0.754103 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} 0.130451 \\ 0.046707 \\ -0.338593 \\ -0.545552 \\ -0.754008 \end{array} \right) & \left(\begin{array}{c} 0.130452 \\ 0.046707 \\ -0.338593 \\ -0.545553 \\ -0.754007 \end{array} \right) &
 \end{array}$$

Matriz de teste 234

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} -0.089750 \\ -0.060340 \\ 0.347183 \\ -0.214015 \\ -0.906623 \end{array} \right) & \left(\begin{array}{c} -0.094300 \\ -0.060272 \\ 0.347164 \\ -0.214331 \\ -0.906098 \end{array} \right) & \left(\begin{array}{c} -0.089750 \\ -0.060340 \\ 0.347183 \\ -0.214015 \\ -0.906623 \end{array} \right) & \left(\begin{array}{c} -0.089572 \\ -0.060344 \\ 0.347177 \\ -0.213940 \\ -0.906661 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} -0.089750 \\ -0.060340 \\ 0.347183 \\ -0.214015 \\ -0.906623 \end{array} \right) & \left(\begin{array}{c} -0.074033 \\ -0.061237 \\ 0.343455 \\ -0.207896 \\ -0.910817 \end{array} \right) & \left(\begin{array}{c} 0.089750 \\ 0.060340 \\ -0.347183 \\ 0.214015 \\ 0.906623 \end{array} \right) & \left(\begin{array}{c} 0.082636 \\ 0.061201 \\ -0.345551 \\ 0.212122 \\ 0.908308 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} -0.089750 \\ -0.060340 \\ 0.347183 \\ -0.214015 \\ -0.906623 \end{array} \right) & \left(\begin{array}{c} -0.087792 \\ -0.060376 \\ 0.345824 \\ -0.212961 \\ -0.907580 \end{array} \right) &
 \end{array}$$

Matriz de teste 240

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} -0.362339 \\ -0.088974 \\ 0.432902 \\ 0.492229 \\ 0.656583 \end{array} \right) & \left(\begin{array}{c} 0.359888 \\ 0.089008 \\ -0.433373 \\ -0.492765 \\ -0.657213 \end{array} \right) & \left(\begin{array}{c} 0.362339 \\ 0.088974 \\ -0.432902 \\ -0.492229 \\ -0.656583 \end{array} \right) & \left(\begin{array}{c} 0.364589 \\ 0.089163 \\ -0.431581 \\ -0.490196 \\ -0.657701 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} -0.362339 \\ -0.088974 \\ 0.432902 \\ 0.492229 \\ 0.656583 \end{array} \right) & \left(\begin{array}{c} 0.361028 \\ 0.091733 \\ -0.436605 \\ -0.492030 \\ -0.654619 \end{array} \right) & \left(\begin{array}{c} -0.362339 \\ -0.088974 \\ 0.432902 \\ 0.492229 \\ 0.656583 \end{array} \right) & \left(\begin{array}{c} 0.361507 \\ 0.087216 \\ -0.432600 \\ -0.493281 \\ -0.656686 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} 0.362339 \\ 0.088974 \\ -0.432902 \\ -0.492229 \\ -0.656583 \end{array} \right) & \left(\begin{array}{c} -0.365272 \\ -0.087963 \\ 0.437665 \\ 0.498863 \\ 0.646857 \end{array} \right) &
 \end{array}$$

Matriz de teste 268

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} 0.603262 \\ 0.205977 \\ 0.157312 \\ 0.171146 \\ -0.734582 \end{array} \right) & \left(\begin{array}{c} 0.602824 \\ 0.206056 \\ 0.157383 \\ 0.171220 \\ -0.734886 \end{array} \right) & \left(\begin{array}{c} -0.603262 \\ -0.205977 \\ -0.157312 \\ -0.171146 \\ 0.734582 \end{array} \right) & \left(\begin{array}{c} -0.603265 \\ -0.205984 \\ -0.157305 \\ -0.171140 \\ 0.734580 \end{array} \right) \\
 \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} 0.603262 \\ 0.205977 \\ 0.157312 \\ 0.171146 \\ -0.734582 \end{array} \right) & \left(\begin{array}{c} -0.601910 \\ -0.203222 \\ -0.160487 \\ -0.173721 \\ 0.735166 \end{array} \right) & \left(\begin{array}{c} -0.603262 \\ -0.205977 \\ -0.157312 \\ -0.171146 \\ 0.734582 \end{array} \right) & \left(\begin{array}{c} 0.603296 \\ 0.206048 \\ 0.157250 \\ 0.171079 \\ -0.734563 \end{array} \right) \\
 \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} 0.603262 \\ 0.205977 \\ 0.157312 \\ 0.171146 \\ -0.734582 \end{array} \right) & \left(\begin{array}{c} -0.603236 \\ -0.205412 \\ -0.157963 \\ -0.171865 \\ 0.734454 \end{array} \right) &
 \end{array}$$

Matriz de teste 287

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} 0.025041 \\ 0.454619 \\ -0.424459 \\ 0.560047 \\ -0.546696 \end{array} \right) & \left(\begin{array}{c} 0.024957 \\ 0.454620 \\ -0.424460 \\ 0.560048 \\ -0.546698 \end{array} \right) & \left(\begin{array}{c} 0.025041 \\ 0.454619 \\ -0.424459 \\ 0.560047 \\ -0.546696 \end{array} \right) & \left(\begin{array}{c} -0.025116 \\ -0.454686 \\ 0.424437 \\ -0.560051 \\ 0.546649 \end{array} \right) \\
 \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} 0.025041 \\ 0.454619 \\ -0.424459 \\ 0.560047 \\ -0.546696 \end{array} \right) & \left(\begin{array}{c} 0.024772 \\ 0.454563 \\ -0.424488 \\ 0.559971 \\ -0.546810 \end{array} \right) & \left(\begin{array}{c} -0.025041 \\ -0.454619 \\ 0.424459 \\ -0.560047 \\ 0.546696 \end{array} \right) & \left(\begin{array}{c} -0.023696 \\ -0.454481 \\ 0.424659 \\ -0.559577 \\ 0.547196 \end{array} \right) \\
 \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} 0.025041 \\ 0.454619 \\ -0.424459 \\ 0.560047 \\ -0.546696 \end{array} \right) & \left(\begin{array}{c} -0.025578 \\ -0.455036 \\ 0.424714 \\ -0.560656 \\ 0.545501 \end{array} \right) &
 \end{array}$$

Apêndice C - Resultados Experimentais

Matrizes 10x10

Matriz de teste 378

m	m_{pert}	M	M_{pert}
$\begin{pmatrix} 0.101170 \\ -0.216583 \\ 0.209534 \\ 0.107198 \\ -0.258452 \\ -0.144862 \\ 0.104773 \\ 0.513342 \\ -0.562213 \\ -0.457271 \end{pmatrix}$	$\begin{pmatrix} 0.101059 \\ -0.216587 \\ 0.209538 \\ 0.107200 \\ -0.258453 \\ -0.144865 \\ 0.104775 \\ 0.513346 \\ -0.562219 \\ -0.457278 \end{pmatrix}$	$\begin{pmatrix} 0.101170 \\ -0.216583 \\ 0.209534 \\ 0.107198 \\ -0.258452 \\ -0.144862 \\ 0.104773 \\ 0.513342 \\ -0.562213 \\ -0.457271 \end{pmatrix}$	$\begin{pmatrix} -0.101179 \\ 0.216548 \\ -0.209563 \\ -0.107206 \\ 0.258467 \\ 0.144899 \\ -0.104762 \\ -0.513321 \\ 0.562199 \\ 0.457295 \end{pmatrix}$
W	W_{pert}	N	N_{pert}
$\begin{pmatrix} 0.101170 \\ -0.216583 \\ 0.209534 \\ 0.107198 \\ -0.258452 \\ -0.144862 \\ 0.104773 \\ 0.513342 \\ -0.562213 \\ -0.457271 \end{pmatrix}$	$\begin{pmatrix} -0.103807 \\ 0.215431 \\ -0.208754 \\ -0.106121 \\ 0.258502 \\ 0.145667 \\ -0.103856 \\ -0.513803 \\ 0.562019 \\ 0.457479 \end{pmatrix}$	$\begin{pmatrix} 0.101170 \\ -0.216583 \\ 0.209534 \\ 0.107198 \\ -0.258452 \\ -0.144862 \\ 0.104773 \\ 0.513342 \\ -0.562213 \\ -0.457271 \end{pmatrix}$	$\begin{pmatrix} 0.100532 \\ -0.216685 \\ 0.209925 \\ 0.107447 \\ -0.258308 \\ -0.144860 \\ 0.104933 \\ 0.513087 \\ -0.562160 \\ -0.457523 \end{pmatrix}$
	A	A_{pert}	
	$\begin{pmatrix} -0.101170 \\ 0.216583 \\ -0.209534 \\ -0.107198 \\ 0.258452 \\ 0.144862 \\ -0.104773 \\ -0.513342 \\ 0.562213 \\ 0.457271 \end{pmatrix}$	$\begin{pmatrix} 0.101603 \\ -0.216727 \\ 0.209344 \\ 0.107087 \\ -0.258664 \\ -0.144810 \\ 0.104749 \\ 0.513842 \\ -0.562607 \\ -0.456074 \end{pmatrix}$	

Matriz de teste 485

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} 0.115926 \\ 0.098329 \\ -0.085886 \\ -0.196866 \\ -0.854505 \\ 0.137798 \\ 0.263748 \\ -0.086722 \\ 0.127058 \\ -0.297264 \end{array} \right) & \left(\begin{array}{c} -0.115903 \\ -0.098330 \\ 0.085886 \\ 0.196867 \\ 0.854507 \\ -0.137798 \\ -0.263749 \\ 0.086722 \\ -0.127058 \\ 0.297264 \end{array} \right) & \left(\begin{array}{c} 0.115926 \\ 0.098329 \\ -0.085886 \\ -0.196866 \\ -0.854505 \\ 0.137798 \\ 0.263748 \\ -0.086722 \\ 0.127058 \\ -0.297264 \end{array} \right) & \left(\begin{array}{c} 0.115926 \\ 0.098329 \\ -0.085886 \\ -0.196867 \\ -0.854505 \\ 0.137798 \\ 0.263749 \\ -0.086720 \\ 0.127057 \\ -0.297264 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} -0.115926 \\ -0.098329 \\ 0.085886 \\ 0.196866 \\ 0.854505 \\ -0.137798 \\ -0.263748 \\ 0.086722 \\ -0.127058 \\ 0.297264 \end{array} \right) & \left(\begin{array}{c} -0.115303 \\ -0.098793 \\ 0.085707 \\ 0.197487 \\ 0.854565 \\ -0.136928 \\ -0.264437 \\ 0.085456 \\ -0.126518 \\ 0.297206 \end{array} \right) & \left(\begin{array}{c} 0.115926 \\ 0.098329 \\ -0.085886 \\ -0.196866 \\ -0.854505 \\ 0.137798 \\ 0.263748 \\ -0.086722 \\ 0.127058 \\ -0.297264 \end{array} \right) & \left(\begin{array}{c} 0.115860 \\ 0.098451 \\ -0.085850 \\ -0.196943 \\ -0.854475 \\ 0.137694 \\ 0.263874 \\ -0.086473 \\ 0.126972 \\ -0.297340 \end{array} \right) \\
 A & A_{pert} & & \\
 \left(\begin{array}{c} 0.115926 \\ 0.098329 \\ -0.085886 \\ -0.196866 \\ -0.854505 \\ 0.137798 \\ 0.263748 \\ -0.086722 \\ 0.127058 \\ -0.297264 \end{array} \right) & \left(\begin{array}{c} -0.116210 \\ -0.098612 \\ 0.086143 \\ 0.197516 \\ 0.856930 \\ -0.138046 \\ -0.264649 \\ 0.086794 \\ -0.127341 \\ 0.288382 \end{array} \right) & &
 \end{array}$$

Matriz de teste 515

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} 0.540931 \\ -0.235576 \\ 0.146399 \\ -0.489688 \\ 0.333212 \\ -0.104968 \\ -0.461974 \\ 0.179305 \\ -0.056508 \\ 0.140920 \end{array} \right) & \left(\begin{array}{c} -0.540933 \\ 0.235575 \\ -0.146399 \\ 0.489688 \\ -0.333211 \\ 0.104967 \\ 0.461973 \\ -0.179305 \\ 0.056508 \\ -0.140920 \end{array} \right) & \left(\begin{array}{c} -0.540931 \\ 0.235576 \\ -0.146399 \\ 0.489688 \\ -0.333212 \\ 0.104968 \\ 0.461974 \\ -0.179305 \\ 0.056508 \\ -0.140920 \end{array} \right) & \left(\begin{array}{c} 0.540931 \\ -0.235575 \\ 0.146399 \\ -0.489689 \\ 0.333212 \\ -0.104968 \\ -0.461974 \\ 0.179305 \\ -0.056508 \\ 0.140920 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} 0.540931 \\ -0.235576 \\ 0.146399 \\ -0.489688 \\ 0.333212 \\ -0.104968 \\ -0.461974 \\ 0.179305 \\ -0.056508 \\ 0.140920 \end{array} \right) & \left(\begin{array}{c} -0.540926 \\ 0.235599 \\ -0.146396 \\ 0.489661 \\ -0.333215 \\ 0.104948 \\ 0.461994 \\ -0.179306 \\ 0.056515 \\ -0.140938 \end{array} \right) & \left(\begin{array}{c} 0.540931 \\ -0.235576 \\ 0.146399 \\ -0.489688 \\ 0.333212 \\ -0.104968 \\ -0.461974 \\ 0.179305 \\ -0.056508 \\ 0.140920 \end{array} \right) & \left(\begin{array}{c} 0.540931 \\ -0.235577 \\ 0.146399 \\ -0.489686 \\ 0.333211 \\ -0.104966 \\ -0.461975 \\ 0.179306 \\ -0.056509 \\ 0.140921 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} 0.540931 \\ -0.235576 \\ 0.146399 \\ -0.489688 \\ 0.333212 \\ -0.104968 \\ -0.461974 \\ 0.179305 \\ -0.056508 \\ 0.140920 \end{array} \right) & \left(\begin{array}{c} -0.540947 \\ 0.235577 \\ -0.146405 \\ 0.489709 \\ -0.333224 \\ 0.104977 \\ 0.461984 \\ -0.179308 \\ 0.056506 \\ -0.140710 \end{array} \right) &
 \end{array}$$

Matriz de teste 528

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} -0.369421 \\ -0.111020 \\ -0.202821 \\ 0.567849 \\ 0.184692 \\ 0.105549 \\ 0.186239 \\ 0.086535 \\ 0.512511 \\ 0.370840 \end{array} \right) & \left(\begin{array}{c} -0.369411 \\ -0.111020 \\ -0.202822 \\ 0.567851 \\ 0.184693 \\ 0.105549 \\ 0.186240 \\ 0.086535 \\ 0.512513 \\ 0.370842 \end{array} \right) & \left(\begin{array}{c} 0.369421 \\ 0.111020 \\ 0.202821 \\ -0.567849 \\ -0.184692 \\ -0.105549 \\ -0.186239 \\ -0.086535 \\ -0.512511 \\ -0.370840 \end{array} \right) & \left(\begin{array}{c} -0.369421 \\ -0.111020 \\ -0.202821 \\ 0.567849 \\ 0.184692 \\ 0.105549 \\ 0.186239 \\ 0.086535 \\ 0.512511 \\ 0.370840 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} -0.369421 \\ -0.111020 \\ -0.202821 \\ 0.567849 \\ 0.184692 \\ 0.105549 \\ 0.186239 \\ 0.086535 \\ 0.512511 \\ 0.370840 \end{array} \right) & \left(\begin{array}{c} 0.369527 \\ 0.110768 \\ 0.202721 \\ -0.567717 \\ -0.185112 \\ -0.105454 \\ -0.185973 \\ -0.086600 \\ -0.512483 \\ -0.371039 \end{array} \right) & \left(\begin{array}{c} 0.369421 \\ 0.111020 \\ 0.202821 \\ -0.567849 \\ -0.184692 \\ -0.105549 \\ -0.186239 \\ -0.086535 \\ -0.512511 \\ -0.370840 \end{array} \right) & \left(\begin{array}{c} 0.369419 \\ 0.111023 \\ 0.202824 \\ -0.567851 \\ -0.184689 \\ -0.105550 \\ -0.186241 \\ -0.086534 \\ -0.512510 \\ -0.370838 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} -0.369421 \\ -0.111020 \\ -0.202821 \\ 0.567849 \\ 0.184692 \\ 0.105549 \\ 0.186239 \\ 0.086535 \\ 0.512511 \\ 0.370840 \end{array} \right) & \left(\begin{array}{c} -0.369406 \\ -0.111035 \\ -0.202833 \\ 0.567855 \\ 0.184673 \\ 0.105555 \\ 0.186250 \\ 0.086527 \\ 0.512502 \\ 0.370849 \end{array} \right) &
 \end{array}$$

Matriz de teste 572

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} 0.098839 \\ -0.097668 \\ 0.235660 \\ 0.451583 \\ -0.034318 \\ 0.216011 \\ 0.262739 \\ 0.708671 \\ -0.301106 \\ -0.107140 \end{array} \right) & \left(\begin{array}{c} -0.098839 \\ 0.097668 \\ -0.235660 \\ -0.451583 \\ 0.034318 \\ -0.216011 \\ -0.262739 \\ -0.708671 \\ 0.301106 \\ 0.107140 \end{array} \right) & \left(\begin{array}{c} -0.098839 \\ 0.097668 \\ -0.235660 \\ -0.451583 \\ 0.034318 \\ -0.216011 \\ -0.262739 \\ -0.708671 \\ 0.301106 \\ 0.107140 \end{array} \right) & \left(\begin{array}{c} -0.098839 \\ 0.097668 \\ -0.235660 \\ -0.451583 \\ 0.034318 \\ -0.216011 \\ -0.262739 \\ -0.708671 \\ 0.301106 \\ 0.107140 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} 0.098839 \\ -0.097668 \\ 0.235660 \\ 0.451583 \\ -0.034318 \\ 0.216011 \\ 0.262739 \\ 0.708671 \\ -0.301106 \\ -0.107140 \end{array} \right) & \left(\begin{array}{c} 0.098840 \\ -0.097669 \\ 0.235657 \\ 0.451582 \\ -0.034319 \\ 0.216009 \\ 0.262738 \\ 0.708673 \\ -0.301107 \\ -0.107138 \end{array} \right) & \left(\begin{array}{c} -0.098839 \\ 0.097668 \\ -0.235660 \\ -0.451583 \\ 0.034318 \\ -0.216011 \\ -0.262739 \\ -0.708671 \\ 0.301106 \\ 0.107140 \end{array} \right) & \left(\begin{array}{c} 0.098839 \\ -0.097668 \\ 0.235660 \\ 0.451583 \\ -0.034318 \\ 0.216011 \\ 0.262739 \\ 0.708671 \\ -0.301106 \\ -0.107140 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} -0.098839 \\ 0.097668 \\ -0.235660 \\ -0.451583 \\ 0.034318 \\ -0.216011 \\ -0.262739 \\ -0.708671 \\ 0.301106 \\ 0.107140 \end{array} \right) & \left(\begin{array}{c} -0.098839 \\ 0.097668 \\ -0.235660 \\ -0.451582 \\ 0.034318 \\ -0.216010 \\ -0.262738 \\ -0.708670 \\ 0.301106 \\ 0.107151 \end{array} \right) &
 \end{array}$$

Matriz de teste 588

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} 0.012081 \\ 0.257512 \\ 0.240581 \\ 0.026019 \\ 0.555024 \\ -0.002951 \\ -0.393364 \\ 0.213121 \\ 0.345037 \\ -0.497713 \end{array} \right) & \left(\begin{array}{c} -0.010045 \\ -0.257740 \\ -0.240236 \\ -0.025795 \\ -0.554893 \\ 0.002839 \\ 0.393170 \\ -0.213219 \\ -0.345338 \\ 0.497868 \end{array} \right) & \left(\begin{array}{c} 0.012081 \\ 0.257512 \\ 0.240581 \\ 0.026019 \\ 0.555024 \\ -0.002951 \\ -0.393364 \\ 0.213121 \\ 0.345037 \\ -0.497713 \end{array} \right) & \left(\begin{array}{c} 0.012168 \\ 0.257354 \\ 0.240681 \\ 0.026201 \\ 0.554956 \\ -0.003009 \\ -0.393544 \\ 0.213256 \\ 0.344794 \\ -0.497779 \end{array} \right) \\
 W & W_{pert} & N & N_{pert} \\
 \left(\begin{array}{c} 0.012081 \\ 0.257512 \\ 0.240581 \\ 0.026019 \\ 0.555024 \\ -0.002951 \\ -0.393364 \\ 0.213121 \\ 0.345037 \\ -0.497713 \end{array} \right) & \left(\begin{array}{c} -0.021432 \\ -0.257113 \\ -0.246678 \\ -0.031963 \\ -0.561056 \\ 0.005511 \\ 0.394022 \\ -0.209219 \\ -0.340430 \\ 0.491760 \end{array} \right) & \left(\begin{array}{c} -0.012081 \\ -0.257512 \\ -0.240581 \\ -0.026019 \\ -0.555024 \\ 0.002951 \\ 0.393364 \\ -0.213121 \\ -0.345037 \\ 0.497713 \end{array} \right) & \left(\begin{array}{c} 0.009881 \\ 0.257837 \\ 0.239109 \\ 0.023747 \\ 0.553926 \\ -0.002275 \\ -0.392670 \\ 0.213261 \\ 0.346961 \\ -0.498793 \end{array} \right) \\
 & A & A_{pert} & \\
 & \left(\begin{array}{c} -0.012081 \\ -0.257512 \\ -0.240581 \\ -0.026019 \\ -0.555024 \\ 0.002951 \\ 0.393364 \\ -0.213121 \\ -0.345037 \\ 0.497713 \end{array} \right) & \left(\begin{array}{c} -0.014478 \\ -0.254774 \\ -0.239610 \\ -0.030378 \\ -0.550612 \\ 0.003758 \\ 0.391413 \\ -0.212828 \\ -0.337892 \\ 0.510575 \end{array} \right) &
 \end{array}$$

Apêndice D - Resultados Experimentais

Matrizes 15x15

Matriz de teste 624

m	m_{pert}	M	M_{pert}
-0.284587	0.284672	0.284587	0.284577
0.562510	-0.562493	-0.562510	-0.562508
-0.260138	0.260133	0.260138	0.260139
0.143840	-0.143837	-0.143840	-0.143846
0.048071	-0.048074	-0.048071	-0.048071
-0.150405	0.150400	0.150405	0.150401
0.180407	-0.180402	-0.180407	-0.180408
-0.286252	0.286246	0.286252	0.286250
0.208435	-0.208429	-0.208435	-0.208446
0.078534	-0.078530	-0.078534	-0.078542
0.339574	-0.339567	-0.339574	-0.339586
-0.040024	0.040021	0.040024	0.040020
-0.151755	0.151750	0.151755	0.151758
0.366984	-0.366974	-0.366984	-0.366983
-0.224897	0.224892	0.224897	0.224886

W	W_{pert}	N	N_{pert}
-0.284587	-0.284368	-0.284587	0.284629
0.562510	0.562653	0.562510	-0.562491
-0.260138	-0.260153	-0.260138	0.260135
0.143840	0.143708	0.143840	-0.143851
0.048071	0.048108	0.048071	-0.048073
-0.150405	-0.150431	-0.150405	0.150410
0.180407	0.180299	0.180407	-0.180411
-0.286252	-0.286159	-0.286252	0.286271
0.208435	0.208901	0.208435	-0.208374
0.078534	0.079014	0.078534	-0.078460
0.339574	0.339413	0.339574	-0.339593
-0.040024	-0.040246	-0.040024	0.039995
-0.151755	-0.151806	-0.151755	0.151746
0.366984	0.366736	0.366984	-0.367007
-0.224897	-0.225039	-0.224897	0.224885

$$\begin{array}{cc}
 A & A_{pert} \\
 \left(\begin{array}{c} 0.284587 \\ -0.562510 \\ 0.260138 \\ -0.143840 \\ -0.048071 \\ 0.150405 \\ -0.180407 \\ 0.286252 \\ -0.208435 \\ -0.078534 \\ -0.339574 \\ 0.040024 \\ 0.151755 \\ -0.366984 \\ 0.224897 \end{array} \right) & \left(\begin{array}{c} 0.284563 \\ -0.562356 \\ 0.260069 \\ -0.143793 \\ -0.048072 \\ 0.150390 \\ -0.180367 \\ 0.286206 \\ -0.208331 \\ -0.078451 \\ -0.339493 \\ 0.040005 \\ 0.151697 \\ -0.366912 \\ 0.225930 \end{array} \right)
 \end{array}$$

Matriz de teste 803

$$\begin{array}{cccc}
 m & m_{pert} & M & M_{pert} \\
 \left(\begin{array}{c} 0.240784 \\ -0.383331 \\ -0.228682 \\ 0.076254 \\ -0.045462 \\ -0.391508 \\ 0.408286 \\ -0.189290 \\ -0.051336 \\ 0.147774 \\ 0.105489 \\ 0.182905 \\ -0.453937 \\ 0.106486 \\ -0.304376 \end{array} \right) & \left(\begin{array}{c} -0.241336 \\ 0.383276 \\ 0.228651 \\ -0.076237 \\ 0.045512 \\ 0.391462 \\ -0.408235 \\ 0.189248 \\ 0.051312 \\ -0.147754 \\ -0.105477 \\ -0.182862 \\ 0.453866 \\ -0.106493 \\ 0.304327 \end{array} \right) & \left(\begin{array}{c} 0.240784 \\ -0.383331 \\ -0.228682 \\ 0.076254 \\ -0.045462 \\ -0.391508 \\ 0.408286 \\ -0.189290 \\ -0.051336 \\ 0.147774 \\ 0.105489 \\ 0.182905 \\ -0.453937 \\ 0.106486 \\ -0.304376 \end{array} \right) & \left(\begin{array}{c} -0.240784 \\ 0.383330 \\ 0.228681 \\ -0.076256 \\ 0.045463 \\ 0.391500 \\ -0.408299 \\ 0.189291 \\ 0.051329 \\ -0.147775 \\ -0.105479 \\ -0.182918 \\ 0.453932 \\ -0.106480 \\ 0.304374 \end{array} \right)
 \end{array}$$

$$\begin{array}{cccc}
 \begin{array}{c} W \\ \left(\begin{array}{c} -0.240784 \\ 0.383331 \\ 0.228682 \\ -0.076254 \\ 0.045462 \\ 0.391508 \\ -0.408286 \\ 0.189290 \\ 0.051336 \\ -0.147774 \\ -0.105489 \\ -0.182905 \\ 0.453937 \\ -0.106486 \\ 0.304376 \end{array} \right) &
 \begin{array}{c} W_{pert} \\ \left(\begin{array}{c} 0.241711 \\ -0.378557 \\ -0.224407 \\ 0.072756 \\ -0.047966 \\ -0.392937 \\ 0.407493 \\ -0.186579 \\ -0.048909 \\ 0.154136 \\ 0.113629 \\ 0.178938 \\ -0.454939 \\ 0.109254 \\ -0.308334 \end{array} \right) &
 \begin{array}{c} N \\ \left(\begin{array}{c} -0.240784 \\ 0.383331 \\ 0.228682 \\ -0.076254 \\ 0.045462 \\ 0.391508 \\ -0.408286 \\ 0.189290 \\ 0.051336 \\ -0.147774 \\ -0.105489 \\ -0.182905 \\ 0.453937 \\ -0.106486 \\ 0.304376 \end{array} \right) &
 \begin{array}{c} N_{pert} \\ \left(\begin{array}{c} -0.240707 \\ 0.384722 \\ 0.229811 \\ -0.077449 \\ 0.046060 \\ 0.391195 \\ -0.408839 \\ 0.189669 \\ 0.051755 \\ -0.145746 \\ -0.103151 \\ -0.183651 \\ 0.453270 \\ -0.105747 \\ 0.303371 \end{array} \right) \\
 \\
 \begin{array}{c} A \\ \left(\begin{array}{c} 0.240784 \\ -0.383331 \\ -0.228682 \\ 0.076254 \\ -0.045462 \\ -0.391508 \\ 0.408286 \\ -0.189290 \\ -0.051336 \\ 0.147774 \\ 0.105489 \\ 0.182905 \\ -0.453937 \\ 0.106486 \\ -0.304376 \end{array} \right) &
 \begin{array}{c} A_{pert} \\ \left(\begin{array}{c} 0.240664 \\ -0.383408 \\ -0.228821 \\ 0.076316 \\ -0.045823 \\ -0.391290 \\ 0.408088 \\ -0.189186 \\ -0.051317 \\ 0.147498 \\ 0.105163 \\ 0.182740 \\ -0.453690 \\ 0.106321 \\ -0.305581 \end{array} \right) & &
 \end{array}
 \end{array}$$

Apêndice E - Resultados Experimentais

Matrizes 20x20

Matriz de teste 1107

m	m_{pert}	M	M_{pert}
0.007219	-0.007194	-0.007219	-0.007218
0.090298	-0.090301	-0.090298	-0.090300
0.317091	-0.317090	-0.317091	-0.317091
-0.021551	0.021549	0.021551	0.021550
-0.265015	0.265014	0.265015	0.265015
0.256272	-0.256272	-0.256272	-0.256271
-0.000380	0.000379	0.000380	0.000380
-0.020381	0.020378	0.020381	0.020379
0.050790	-0.050790	-0.050790	-0.050789
0.361108	-0.361109	-0.361108	-0.361109
0.027598	-0.027602	-0.027598	-0.027600
-0.109883	0.109880	0.109883	0.109882
0.170741	-0.170741	-0.170741	-0.170741
0.223329	-0.223329	-0.223329	-0.223328
-0.275096	0.275096	0.275096	0.275096
0.042743	-0.042743	-0.042743	-0.042743
-0.155186	0.155185	0.155186	0.155185
0.070428	-0.070427	-0.070428	-0.070428
-0.492356	0.492356	0.492356	0.492356
0.425035	-0.425036	-0.425035	-0.425035

W	W_{pert}	N	N_{pert}
$\left(\begin{array}{c} -0.007219 \\ -0.090298 \\ -0.317091 \\ 0.021551 \\ 0.265015 \\ -0.256272 \\ 0.000380 \\ 0.020381 \\ -0.050790 \\ -0.361108 \\ -0.027598 \\ 0.109883 \\ -0.170741 \\ -0.223329 \\ 0.275096 \\ -0.042743 \\ 0.155186 \\ -0.070428 \\ 0.492356 \\ -0.425035 \end{array} \right)$	$\left(\begin{array}{c} -0.004654 \\ -0.097346 \\ -0.309792 \\ 0.018169 \\ 0.268900 \\ -0.254278 \\ -0.008650 \\ 0.019458 \\ -0.054016 \\ -0.353610 \\ -0.023732 \\ 0.103172 \\ -0.170246 \\ -0.230794 \\ 0.278055 \\ -0.038910 \\ 0.160417 \\ -0.074626 \\ 0.491737 \\ -0.428290 \end{array} \right)$	$\left(\begin{array}{c} 0.007219 \\ 0.090298 \\ 0.317091 \\ -0.021551 \\ -0.265015 \\ 0.256272 \\ -0.000380 \\ -0.020381 \\ 0.050790 \\ 0.361108 \\ 0.027598 \\ -0.109883 \\ 0.170741 \\ 0.223329 \\ -0.275096 \\ 0.042743 \\ -0.155186 \\ 0.070428 \\ -0.492356 \\ 0.425035 \end{array} \right)$	$\left(\begin{array}{c} -0.007669 \\ -0.089181 \\ -0.318164 \\ 0.022101 \\ 0.264529 \\ -0.256591 \\ 0.001511 \\ 0.020574 \\ -0.050338 \\ -0.362008 \\ -0.027934 \\ 0.110927 \\ -0.170793 \\ -0.222353 \\ 0.274639 \\ -0.043173 \\ 0.154521 \\ -0.069867 \\ 0.492391 \\ -0.424563 \end{array} \right)$
	A	A_{pert}	
	$\left(\begin{array}{c} -0.007219 \\ -0.090298 \\ -0.317091 \\ 0.021551 \\ 0.265015 \\ -0.256272 \\ 0.000380 \\ 0.020381 \\ -0.050790 \\ -0.361108 \\ -0.027598 \\ 0.109883 \\ -0.170741 \\ -0.223329 \\ 0.275096 \\ -0.042743 \\ 0.155186 \\ -0.070428 \\ 0.492356 \\ -0.425035 \end{array} \right)$	$\left(\begin{array}{c} -0.007457 \\ -0.089710 \\ -0.317666 \\ 0.021844 \\ 0.264767 \\ -0.256450 \\ 0.000978 \\ 0.020485 \\ -0.050555 \\ -0.361595 \\ -0.027775 \\ 0.110440 \\ -0.170775 \\ -0.222822 \\ 0.274863 \\ -0.042972 \\ 0.154840 \\ -0.070135 \\ 0.492388 \\ -0.424728 \end{array} \right)$	