Entradas diagonais, valores singulares e convexidade

Matheus Brioschi Herkenhoff Vieira

Mestrado em Matemática

Universidade Federal do Espírito Santo

Vitória
Abril 2009

Entradas diagonais, valores singulares e convexidade

Matheus Brioschi Herkenhoff Vieira

Dissertação submetida ao Programa de Pós-Graduação em Matemática da Universidade Federal do Espírito Santo como requisito parcial para a obtenção do grau de Mestre em Matemática.

Aprovada em 03/04/2009 por:

Prof. Ricardo Soares Leite - Orientador, UFES
Prof. Carlos Tomei, PUC-Rio
Prof. José Miguel Malacarne, UFES

Vitória
Abril 2009

Vieira, Matheus B. H., 1982
Entradas diagonais, valores singulares e convexidade. [Vitória] 2009
xii, 43p., 29,7 cm (UFES, M. Sc., Matemática, 2009)
Dissertação, Universidade Federal do Espírito Santo, PPGMAT.
I. Análise
I. PPGMAT/UFES II. Título

Agradecimentos

Aos meus pais Fernando e Liliane, ao meu irmão Bruno, a toda minha família e especialmente à Silvia pelo carinho e apoio incondicional.

Aos amigos.

Ao meu orientador Ricardo Soares Leite.

Aos professores da banca José Miguel Malacarne e Carlos Tomei.

À FAPES pela bolsa de mestrado.

Resumo

Apresentamos uma nova prova do Teorema de Thompson devida a R. S. Leite usando técnicas semelhantes as de um artigo anterior de R. S. Leite, C. Tomei. e T. R. W. Richa. Dados $\sigma_1 > \ldots > \sigma_n > 0$, denotemos por $\mathcal{P}$ o fecho convexo dos pontos $(\pm\sigma_{\pi(1)}, \ldots, \pm\sigma_{\pi(n)})$ tendo um número par de sinais negativos, onde π é uma permutação. O Teorema de Thompson afirma que a diagonal de uma matriz real $n \times n$ com determinante positivo e valores singulares $\sigma_1, \ldots, \sigma_n$ está em $\mathcal{P}$, e dado um vetor d em $\mathcal{P}$, existe uma matriz real $n \times n$ com determinante positivo, valores singulares $\sigma_1, \ldots, \sigma_n$ e diagonal d.

Abstract

We present a new proof of Thompson theorem due to R. S. Leite using similar techniques from a previous paper of R. S. Leite, C. Tomei and T. R. W. Richa. Given $\sigma_1 > \ldots > \sigma_n > 0$, denote by $\mathcal{P}$ the convex hull of the points $(\pm\sigma_{\pi(1)}, \ldots, \pm\sigma_{\pi(n)})$ having a even number of negative signs, where π is a permutation. The Thompson theorem asserts that the diagonal of a real $n \times n$ matrix with positive determinant and singular values $\sigma_1, \ldots, \sigma_n$ is in $\mathcal{P}$, and given a vector d in $\mathcal{P}$, there is a real $n \times n$ matrix with positive determinant, singular values $\sigma_1, \ldots, \sigma_n$ and diagonal d.

Sumário

Notações

$\mathcal{M}(n) = \{\text{matrizes reais } n \times n\}$
$\mathcal{O}(n) = \{\text{matrizes ortogonais reais } n \times n\}$
$\mathcal{SO}(n) = \{\text{matrizes ortogonais reais } n \times n \text{ com determinante igual a } 1\}$
$\mathcal{S}(n) = \{\text{matrizes simétricas reais } n \times n\}$
$\mathcal{A}(n) = \{\text{matrizes anti-simétricas reais } n \times n\}$
$\mathcal{S}'(n) = \{\text{matrizes simétricas reais } n \times n \text{ com espectro simples}\}$
$I_n =$ matriz identidade $n \times n$
$e_i =$ i-ésimo vetor canônico de $\mathbb{R}^n$
$M^T =$ transposta da matriz M
$\text{diag}(M) = (m_{11}, \ldots, m_{nn})$
$\text{tr}(M) = m_{11} + \ldots + m_{nn}$
$\lambda(M) = (\lambda_1, \ldots, \lambda_n),$ onde $\lambda_1 \leq \ldots \leq \lambda_n$ são os autovalores de M
$\text{int}(X) =$ interior de X
$\text{fr}(X) =$ fronteira de X
$\text{ext}(X) =$ exterior de X
$X^\perp =$ complemento ortogonal de X.

Capítulo 1

Introdução

Existem vários teoremas que afirmam que as entradas diagonais das matrizes de uma certa variedade formam um politopo convexo. Talvez o mais conhecido destes teoremas seja o Teorema de Schur-Horn. Para enunciá-lo, consideremos Λ uma matriz real simétrica $n \times n$ com autovalores $\lambda_1, \ldots, \lambda_n$, e denotemos por $\mathcal{A}_\Lambda \subset \mathbb{R}^n$ o fecho convexo dos pontos da forma $(\lambda_{\pi(1)}, \ldots, \lambda_{\pi(n)})$, onde π é uma permutação.

Teorema de Schur-Horn
A diagonal de uma matriz real simétrica $n \times n$ com autovalores $\lambda_1, \ldots, \lambda_n$ está em $\mathcal{A}_\Lambda$. Dado um vetor d em $\mathcal{A}_\Lambda$ existe uma matriz real simétrica $n \times n$ com diagonal d e autovalores $\lambda_1, \ldots, \lambda_n$.

O conjunto de todas as matrizes reais simétricas $n \times n$ com autovalores $\lambda_1, \ldots, \lambda_n$ é uma variedade real, que denotamos por O_Λ. Podemos também enunciar o Teorema de Schur-Horn, de maneira mais conveniente para as técnicas utilizadas nesta dissertação, do seguinte modo:

Teorema de Schur-Horn
A imagem da função

$$\begin{aligned} SH : O_\Lambda &\to \mathbb{R}^n \\ S &\mapsto \operatorname{diag}(S) \end{aligned}$$

é o politopo $\mathcal{A}_\Lambda$.

Todas as matrizes reais simétricas $n \times n$ com autovalores $\lambda_1, \ldots, \lambda_n$ têm o mesmo traço: $\lambda_1 + \ldots + \lambda_n$. Logo podemos projetar o politopo $\mathcal{A}_\Lambda$ no hiperplano do $\mathbb{R}^n$ dado por $\sum_{i=1}^n x_i = \sum_{i=1}^n \lambda_i$. As Figuras 1.1 e 1.2 mostram o politopo $\mathcal{A}_\Lambda$ para Λ com autovalores $4, 2, 1$ e autovalores $7, 4, 2, 1$, respectivamente.

A primeira parte do Teorema de Schur-Horn, que diz que a imagem da função SH está contida no politopo, foi apresentada por Schur [5] em 1923. A segunda inclusão, que todo ponto do politopo $\mathcal{A}_\Lambda$ é a imagem de uma matriz de O_Λ, é devida a Horn [1] em 1954. Este resultado motivou o teorema de convexidade de Kostant para álgebras de Lie [3] em 1973, que implica a versão complexa do Teorema de Schur-Horn. Sucessivas generalizações culminaram com o teorema da convexidade da imagem aplicação momento, obtido por Atiyah [7] e Guillemin e Sternberg [13] de maneira independente em 1982. Usando este resultado, em 1988, Duistermaat [6] deu uma nova demonstração do caso real do Teorema de Schur-Horn.

No mesmo artigo em que apresenta a demonstração da segunda inclusão, Horn demonstra que as entradas diagonais de matrizes em $\mathcal{SO}(n)$, matrizes ortogonais reais $n \times n$ com determinante positivo, também formam um politopo convexo.

Teorema de Horn para matrizes ortogonais
A imagem da função

$$\begin{aligned} H : \mathcal{SO}(n) &\to \mathbb{R}^n \\ Q &\mapsto \operatorname{diag}(Q) \end{aligned}$$

é o politopo dado pelo fecho convexo dos pontos da forma $(\pm 1, \ldots, \pm 1)$ tendo um número par de coordenadas iguais a -1.

A Figura 1.3 mostra o politopo de Horn em dimensão 3.

Agora enunciamos a versão real do Teorema de Thompson para valores singulares.

Seja Σ uma matriz real diagonal $n \times n$ com entradas diagonais estritamente positivas $\sigma_1 > \ldots > \sigma_n > 0$. Denotemos por $\mathcal{P}$ o politopo dado pelo fecho convexo dos pontos $(\pm\sigma_{\pi(1)}, \ldots, \pm\sigma_{\pi(n)})$ tendo um número par de entradas -1, onde π é uma permutação.

Teorema de Thompson
A diagonal de uma matriz real $n \times n$ com determinante positivo e valores singulares $\sigma_1 > \ldots > \sigma_n > 0$ está em $\mathcal{P}$. Dado um vetor d em $\mathcal{P}$, existe uma matriz real $n \times n$ com determinante positivo, diagonal d e valores singulares $\sigma_1 > \ldots > \sigma_n > 0$.

Denotemos por O_Σ^+ o conjunto das matrizes com determinante positivo e valores singulares $\sigma_1 > \ldots > \sigma_n > 0$. Temos que O_Σ^+ é o conjunto das matrizes de determinante positivo da forma $U\Sigma V$, onde U e V são matrizes ortogonais reais $n \times n$. Podemos enunciar o Teorema de Thompson da seguinte forma:

Teorema de Thompson
A imagem da função

$$\begin{aligned} F : O_\Sigma^+ &\to \mathbb{R}^n \\ M &\mapsto \operatorname{diag}(M) \end{aligned}$$

é o politopo $\mathcal{P}$.

As Figuras 1.4 e 1.5 são exemplos do politopo de Thompson para Σ com diagonal $(2,1)$ e $(4,2,1)$.

Em 1977 Thompson [11] deu a primeira demonstração do teorema que ganhou seu nome para os casos real e complexo. Tam [12] deu uma nova prova para ambos os casos com técnicas simpléticas em 1999.

Em 1999, Leite, Richa e Tomei [10] apresentaram uma nova prova para o Teorema de Schur-Horn, o Teorema de Horn [1], e enunciaram e provaram o que poderia ser chamado de versão anti-simétrica de Schur-Horn. Os três teoremas consideram funções em variedades de matrizes cujos valores são certas entradas destas matrizes e afirmam que a imagem destas funções são politopos. Para o Teorema de Schur-Horn, a variedade em questão é O_Λ e a função é SH. Para Horn, a variedade é $\mathcal{SO}(n)$ e a função é H. Para a versão anti-simétrica de Schur-Horn a variedade é O_Γ e a função é $SHAS$, que serão descritas logo abaixo. A demonstração é semelhante em todos os casos. Mais adiante falaremos mais sobre a versão anti-simétrica do Teorema de Schur-Horn.

A ferramenta usada no artigo acima citado para a prova do Teorema de Schur-Horn foi cálculo em variedades, isso fez com que se utilizasse uma descrição para o politopo através de desigualdades. Explicando melhor, o politopo P_Λ que foi definido acima pode ser visto com a solução de um sistema de inequações lineares. Analiticamente é mais interessante trabalhar com desigualdades do que com fecho convexo. Estas idéias também foram aplicadas no Teorema de Horn e na versão anti-simétrica de Schur-Horn.

Nesse texto demonstraremos o caso real do Teorema de Thompson utilizando cálculo em variedades, seguindo as idéias do artigo [10]. Esta prova é devida a R. S. Leite. Naturalmente, usaremos a descrição analiticamente mais conveniente para o politopo do Teorema, a saber, a que o expressa através de um sistema de desigualdades. Esta descrição somente será feita quando necessária, pois a forma padrão de definirmos o politopo do Teorema de Thompson é mais elegante.

Existe uma versão do Teorema de Thompson para matrizes de determinante negativo.

Teorema de Thompson
A diagonal de uma matriz real $n \times n$ com determinante negativo e valores singulares $\sigma_1 > \ldots > \sigma_n > 0$ está no fecho convexo dos pontos $(\pm\sigma_{\pi(1)}, \ldots, \pm\sigma_{\pi(n)})$ tendo um número ímpar de entradas -1, onde π é uma permutação. Dado um vetor d no fecho convexo dos pontos $(\pm\sigma_{\pi(1)}, \ldots, \pm\sigma_{\pi(n)})$ tendo um número ímpar de entradas -1, onde π é uma permutação, existe uma matriz real $n \times n$ com determinante negativo, diagonal d e valores singulares $\sigma_1 > \ldots > \sigma_n > 0$.

As Figuras 1.6 e 1.7 são exemplos do politopo de Thompson para matrizes de determinante negativo, correspondentes às Figuras 1.4 e 1.5. As Figuras 1.8

e 1.9 são os politopos das versões positiva e negativa do Teorema de Thompson numa só figura.

Para enunciar a versão real anti-simétrica do Teorema de Schur-Horn, precisamos de algumas definições.

Dadas uma matriz Γ da forma

$$\begin{pmatrix} 0 & \Lambda \\ -\Lambda & 0 \end{pmatrix} \text{ ou } \begin{pmatrix} 0 & 0 & \Lambda \\ 0 & 0 & 0 \\ -\Lambda & 0 & 0 \end{pmatrix}$$

onde Λ é uma matriz diagonal $n \times n$, denotamos por O_Γ o conjunto das matrizes da forma $Q\Gamma Q^T$ com Q em $\mathcal{SO}(2n)$ ou $\mathcal{SO}(2n+1)$ dependendo de qual das formas acima é Γ. Denotamos por D_Γ o politopo dado pelo fecho convexo dos pontos $(\pm\sigma_{\pi(1)}, \ldots, \pm\sigma_{\pi(n)})$ tendo um número par de entradas -1, onde π é uma permutação e por E_Γ o politopo dado pelo fecho convexo dos pontos $(\pm\sigma_{\pi(1)}, \ldots, \pm\sigma_{\pi(n)})$.

Definimos a subdiagonal de uma matriz como sendo

$$M \to (m_{1,n+1}, \ldots, m_{n,2n}) \text{ ou } M \to (m_{1,n+2}, \ldots, m_{n,2n+1})$$

de acordo com a ordem da matriz ser $2n$ ou $2n+1$.

Teorema de Schur-Horn anti-simétrico
A imagem da função

$$\begin{aligned} SHAS : O_\Gamma &\to \mathbb{R}^n \\ M &\mapsto \operatorname{subdiag}(M) \end{aligned}$$

é o politopo D_Γ quando a dimensão é par e E_Γ quando a dimensão é ímpar.

Agora esboçaremos as etapas da prova do Teorema de Thompson.

Primeiro provaremos que a imagem de O_Σ^+ pela função diagonal está contida no politopo $\mathcal{P}$. Começaremos caracterizando os pontos críticos de F: a restrição da função diagonal a O_Σ^+. Nesse momento aparece de forma natural uma família finita de funções $F_D : O_\Sigma^+ \to \mathbb{R}$, onde D é uma matriz diagonal, dadas por $F_D(M) = \operatorname{tr}(MD)$ com a seguinte propriedade: um ponto de O_Σ^+ é ponto crítico de F se e somente é ponto crítico de alguma F_D. Depois descreveremos $\mathcal{P}$ como o conjunto solução de um número finito de inequações da forma $a_\alpha \cdot x \leq b_\alpha$, onde $a_\alpha, x \in \mathbb{R}^n$ e $b \in \mathbb{R}$. Usando essa descrição para o politopo $\mathcal{P}$, juntamente com a equivalência dos pontos críticos de F com a família de funções F_D, e observando que um ponto $M \in O_\Sigma^+$ satisfaz $\operatorname{diag}(A_\alpha) \cdot F(M) \leq b_\alpha$ se e somente se satisfaz $F_{A_\alpha}(M) \leq b_\alpha$, através de estimativas de máximos e mínimos para as F_D, concluiremos esta primeira parte.

Depois mostraremos que o politopo $\mathcal{P}$ está contido na imagem de O_Σ^+ pela função diagonal. Primeiro veremos que a fronteira de $F(O_\Sigma^+)$ está contida na união de uma família finita de conjuntos chamados hiperplanos críticos. Fazendo um estudo da segunda derivada das funções F_D, segue que a fronteira de $F(O_\Sigma^+)$ está contida na fronteira do politopo $\mathcal{P}$, com exceção de um conjunto insignificante de seus pontos. Usando argumentos de densidade e conexidade, concluiremos esta parte.

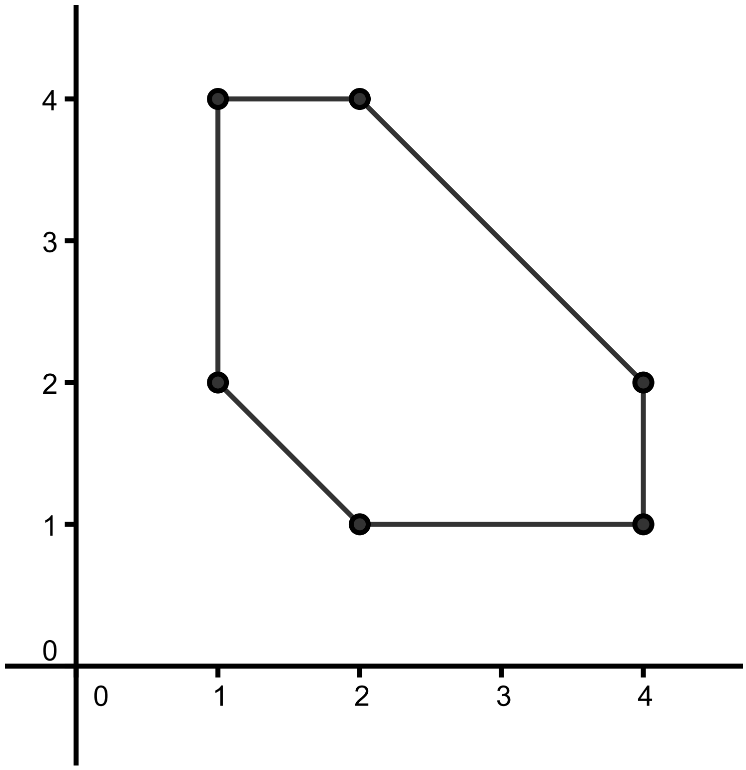

Figura 1.1: O politopo $\mathcal{A}_\Lambda$ para $\Lambda = (4, 2, 1)$ projetado no plano $x + y + z = 7$.

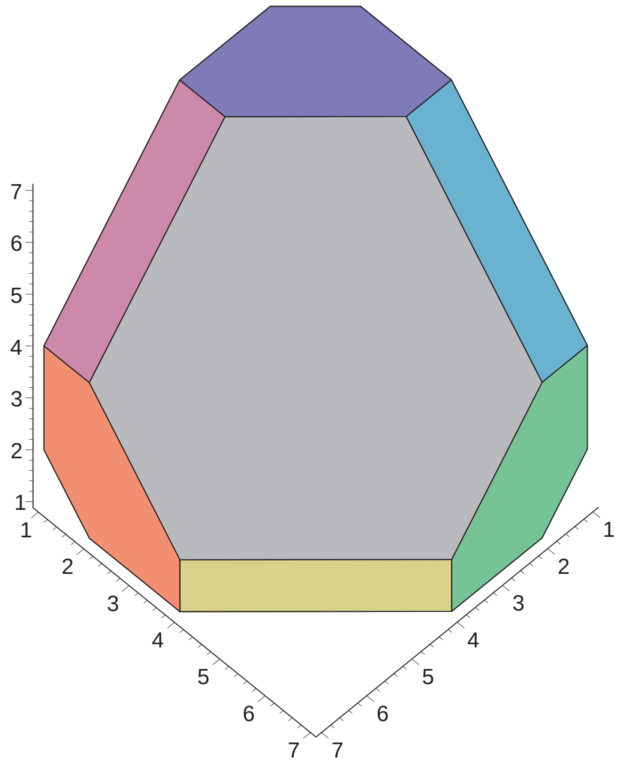

Figura 1.2: O politopo $\mathcal{A}_\Lambda$ para $\Lambda = (7, 4, 2, 1)$ projetado no plano $x_1 + x_2 + x_3 + x_4 = 14$.

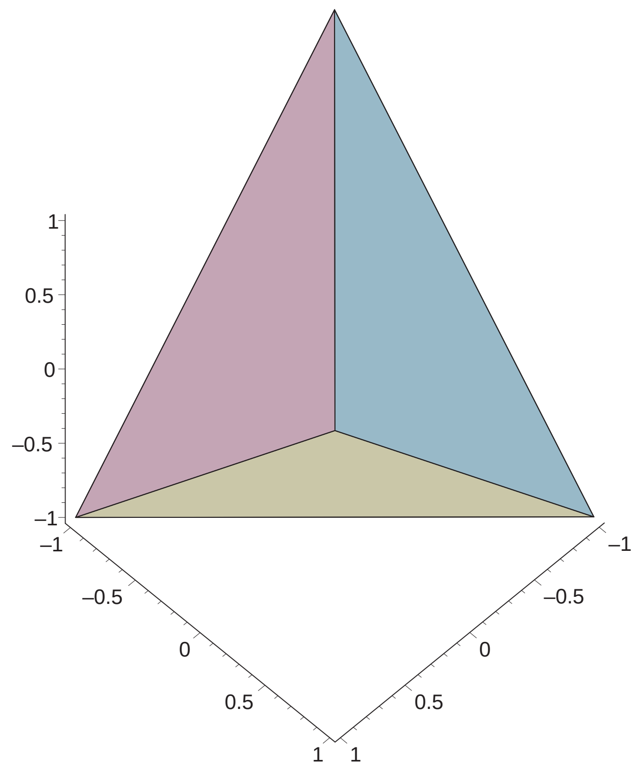

Figura 1.3: O politopo de Horn em dimensão 3.

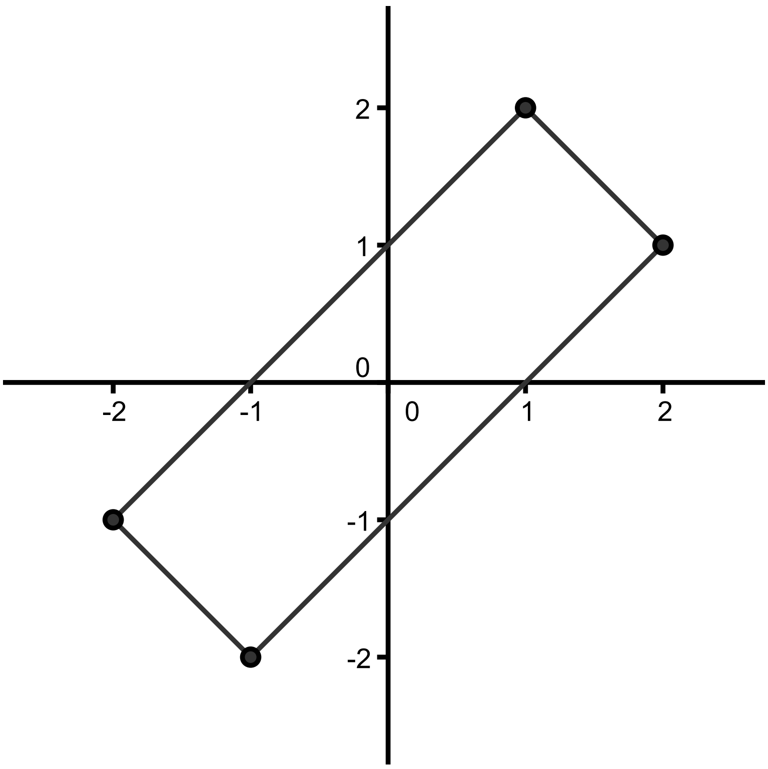

Figura 1.4: O politopo $\mathcal{P}$ para $\Sigma = (2, 1)$.

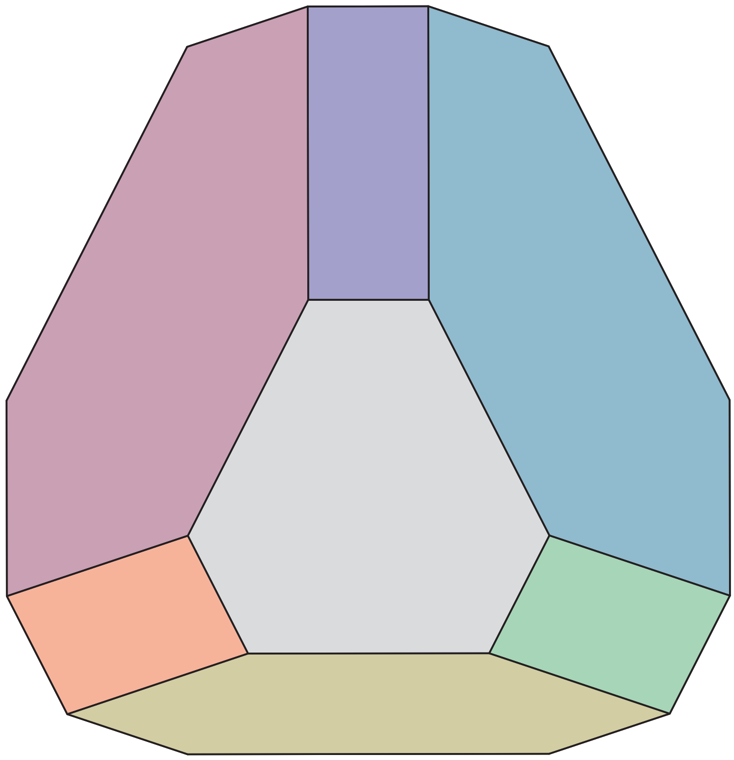

Figura 1.5: O politopo $\mathcal{P}$ para $\Sigma = (4, 2, 1)$.

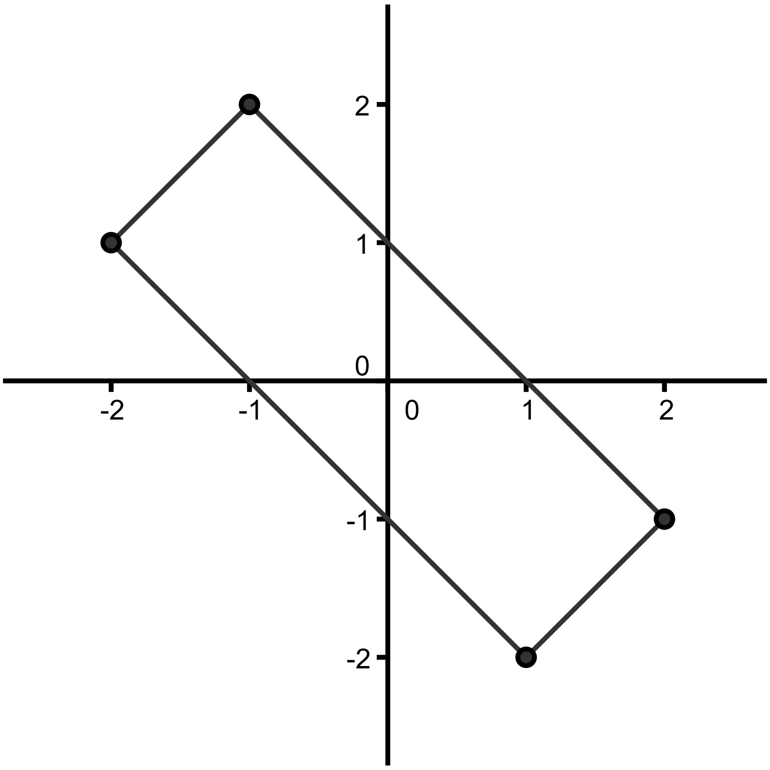

Figura 1.6: O politopo de Thompson (versão determinante negativo) para $\Sigma = (2, 1)$.

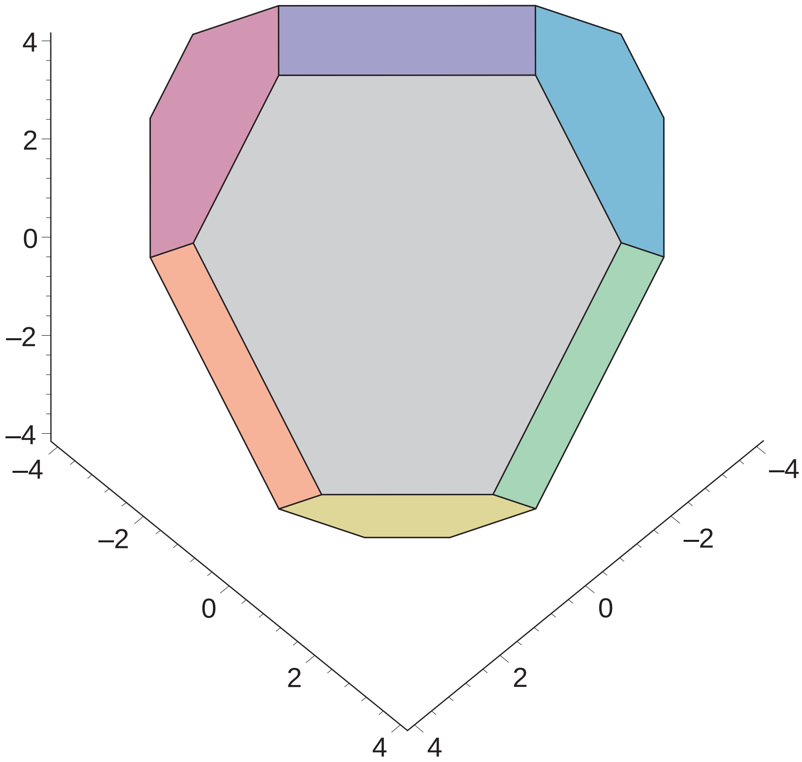

Figura 1.7: O politopo de Thompson (versão determinante negativo) para $\Sigma = (4, 2, 1)$.

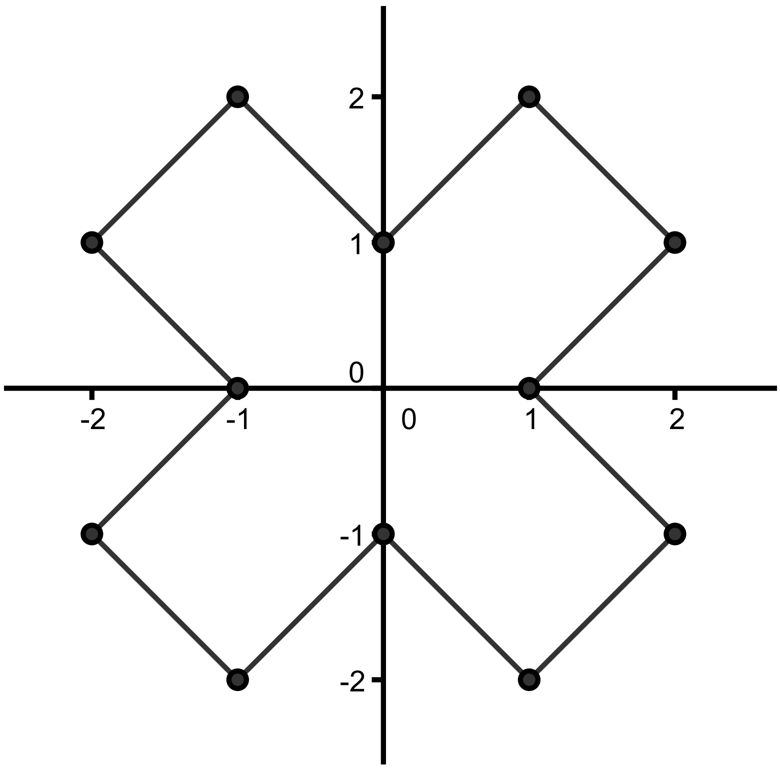

Figura 1.8: O politopo de Thompson (versão determinante negativo e positivo) para $\Sigma = (2, 1)$.

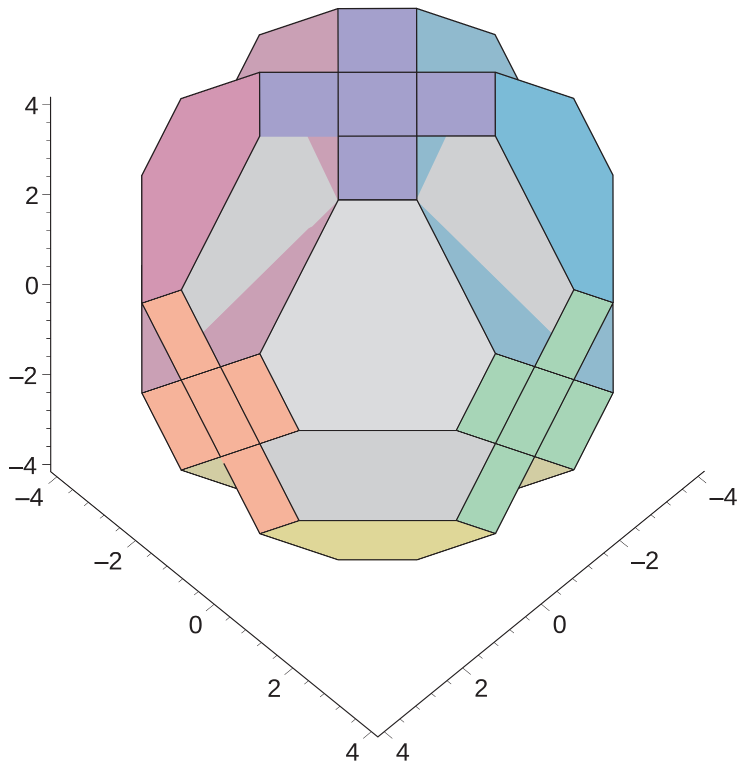

Figura 1.9: O politopo de Thompson (versão determinante negativo e positivo) para $\Sigma = (4, 2, 1)$.

Capítulo 2

A variedade O_Σ^+

Dada uma matriz diagonal Σ com entradas diagonais positivas e estritamente decrescentes, ou seja, $\text{diag}(\Sigma) = (\sigma_1, \ldots, \sigma_n)$ e $\sigma_1 > \ldots > \sigma_n > 0$, lembramos que O_Σ é o conjunto das matrizes da forma $U\Sigma V$, onde U e V são matrizes ortogonais, e que O_Σ^+ é formado pelas matrizes de O_Σ que possuem determinante positivo. Podemos ver O_Σ^+ também como o conjunto das matrizes da forma $U\Sigma V$, onde $U, V \in \mathcal{SO}(n)$.

O objetivo deste capítulo é mostrar o seguinte teorema.

Teorema 1 *O conjunto O_Σ^+ é uma variedade compacta e conexa de dimensão $n^2 - n$. Uma carta em torno do ponto $M \in O_\Sigma^+$ é dada pela função $\psi_M(A, B) = \exp(A)M\exp(B)$, sendo esta definida sobre um aberto de $\mathcal{A}(n) \times \mathcal{A}(n)$. O espaço tangente no ponto M é $T(O_\Sigma^+)_M = \{AM + MB : A, B \in \mathcal{A}(n)\}$.*

Provaremos este teorema através de quatro lemas.

Lema 2 *O conjunto O_Σ^+ é compacto.*

Prova: Primeiro mostraremos que O_Σ é compacto. Para vermos que O_Σ é fechado, definimos a função

$$\begin{aligned} f :& \mathcal{M}(n) \to \mathbb{R}^n \\ & M \mapsto (\det(MM^T - \sigma_1^2 I_n), \ldots, \det(MM^T - \sigma_n^2 I_n)). \end{aligned}$$

É claro que $f^{-1}(0, \ldots, 0) = O_\Sigma$ e f é contínua, portanto O_Σ é fechado. Para provarmos que O_Σ é limitado, consideramos a norma em $\mathcal{M}(n)$ associada ao produto interno $\langle M, N\rangle = \text{tr}(MN^T)$. Dado $M = U\Sigma V \in O_\Sigma$, temos que $|U\Sigma V|^2 = \text{tr}(U\Sigma V(U\Sigma V)^T) = \text{tr}(U\Sigma^2 U^T) = \text{tr}(\Sigma^2) = \text{tr}(\Sigma\Sigma^T) = |\Sigma|^2$, logo O_Σ é limitado.

Agora mostraremos que O_Σ^+ é compacto. Lembrando que o determinante de matrizes ortogonais é sempre ± 1, dado $M = U\Sigma V \in O_\Sigma$, temos que $\det(M) = \det(U\Sigma V) = \pm\det(\Sigma)$. Denotando por D a restrição da função determinante ao conjunto O_Σ, fica claro que $O_\Sigma^+ = D^{-1}(\det(\Sigma))$. Como D é contínua, segue que O_Σ^+ é um conjunto fechado contido no conjunto compacto O_Σ, e portanto é compacto. ■

Lema 3 *O conjunto O_Σ^+ é conexo.*

Prova: Definindo a função

$$\begin{aligned} g :& \mathcal{SO}(n) \times \mathcal{SO}(n) \to O_\Sigma^+ \\ &(U, V) \mapsto U\Sigma V, \end{aligned}$$

e usando que $\mathcal{SO}(n)$ é conexo, g é contínua e $O_\Sigma^+ = g(\mathcal{SO}(n) \times \mathcal{SO}(n))$, concluímos que O_Σ^+ é conexo. ■

Lema 4 *O conjunto O_Σ^+ é uma variedade de dimensão $n^2 - n$.*

Prova: Lembramos que se uma matriz M tem n autovalores $\lambda_1 \leq \ldots \leq \lambda_n$, o espectro ordenado de M é $\lambda_o(M) = (\lambda_1, \ldots, \lambda_n)$. O Teorema espectral garante que λ_o é definida para qualquer matriz simétrica. É claro que para qualquer matriz M, a matriz MM^T é simétrica. Definindo as funções

$$\begin{aligned} g :& \mathcal{M}(n) \to \mathcal{S}(n) \\ &M \mapsto MM^T \end{aligned}$$

e $f = \lambda_o \circ g$, temos que $O_\Sigma = f^{-1}(\sigma_1^2, \ldots, \sigma_n^2)$ e $O_\Sigma^+ = f^{-1}(\sigma_1^2, \ldots, \sigma_n^2) \cap \{$ matrizes de determinante positivo $\}$.

Primeiro mostraremos que f é suave num aberto contendo O_Σ. É fácil ver que g é suave, portanto basta provarmos que λ_o é suave num aberto contendo $g(O_\Sigma)$. Pelo Lema 22 do Apêndice, o conjunto $\mathcal{S}'(n) = \{$ matrizes simétricas $n \times n$ com espectro simples $\}$ é aberto em $\mathcal{S}(n)$, e a restrição da função λ_o ao conjunto $\mathcal{S}'(n) \subset \mathcal{S}(n)$ é suave. Como $g(O_\Sigma) = \{U\Sigma^2U^T : U \in \mathcal{O}(n)\} \subset S'(n)$, segue a afirmação.

Agora veremos que O_Σ é uma variedade de dimensão $n^2 - n$. Pelo Lema 19 do Apêndice, é suficiente verificarmos que

$$df_M : \mathcal{M}(n) \to \mathbb{R}^n$$

é sobrejetiva para todo $M \in O_\Sigma$. Pela regra da cadeia, temos que $df_M = d\lambda_{g(M)} dg_M$, portanto só precisamos provar que

$$dg_M : \mathcal{M}(n) \to \mathcal{S}(n) \quad \text{e} \quad d(\lambda_o)_{g(M)} : \mathcal{S}(n) \to \mathbb{R}^n$$

são sobrejetivas para todo $M \in O_\Sigma$.

Vejamos que dg_M é sobrejetiva para todo $M \in O_\Sigma$. Dados $M, N \in \mathcal{M}(n)$, temos que

$$dg_M(N) = \lim_{t\to 0} \frac{(M+tN)(M+tN)^T - MM^T}{t} = MN^T + NM^T.$$

Fixados $M = U\Sigma V \in O_\Sigma$ e $S \in \mathcal{S}(n)$, encontraremos $N \in \mathcal{M}(n)$ tal que $dg_M(N) = S$. Fazendo $N = \frac{1}{2}SU\Sigma^{-1}V$, temos que

$$\begin{aligned} dg_M(N) &= U\Sigma V(\frac{1}{2}SU\Sigma^{-1}V)^T + (\frac{1}{2}SU\Sigma^{-1}V)(U\Sigma V)^T \\ &= \frac{1}{2}((U\Sigma VV^T\Sigma^{-1}U^T)S^T + S(U\Sigma^{-1}VV^T\Sigma U^T)) = \frac{1}{2}S + \frac{1}{2}S = S. \end{aligned}$$

Portanto dg_M é sobrejetiva para todo $M \in O_\Sigma$.

Agora provaremos que $d(\lambda_o)_{g(M)}$ é sobrejetiva para todo $M \in O_\Sigma$. Dados $M = U\Sigma V \in O_\Sigma$ e $1 \le i \le n$, encontraremos $S \in \mathcal{S}(n)$ tal que $d(\lambda_o)_{g(M)}(S) = e_i$. Fazendo $N = UE_iU^T$, onde E_i é uma matriz diagonal com $\operatorname{diag}(E_i) = e_i$, obtemos

$$\begin{aligned} d(\lambda_o)_{g(M)}(N) &= \lim_{t\to 0} \frac{\lambda_o(g(M)+tN) - \lambda_o(g(M))}{t} \\ &= \lim_{t\to 0} \frac{\lambda_o(U\Sigma^2U^T + tUE_iU^T) - \lambda_o(U\Sigma^2U^T)}{t} \\ &= \lim_{t\to 0} \frac{\lambda_o(U(\Sigma^2 + tE_i)U^T) - \lambda_o(U\Sigma^2U^T)}{t} \\ &= \lim_{t\to 0} \frac{\lambda_o(\Sigma^2 + tE_i) - \lambda_o(\Sigma^2)}{t} = \lim_{t\to 0} \frac{te_i}{t} = e_i. \end{aligned}$$

Portanto $d(\lambda_o)_{g(M)}$ é sobrejetiva para todo $M \in O_\Sigma$. Concluímos que O_Σ é uma variedade de dimensão $n^2 - n$.

Finalmente mostraremos que O_Σ^+ é uma variedade de dimensão $n^2 - n$. Definindo O_Σ^- como o conjunto das matrizes de O_Σ com determinante negativo, temos que $O_\Sigma = O_\Sigma^+ \cup O_\Sigma^-$ e $O_\Sigma^+ \cap O_\Sigma^- = \emptyset$. Como no Lema acima, segue que O_Σ^- é fechado em O_Σ, portanto O_Σ^+ é aberto em O_Σ. Logo O_Σ^+ é uma variedade de dimensão $n^2 - n$. ■

Lema 5 *Uma carta em torno do ponto $M \in O_\Sigma^+$ é dada pela função $\psi_M(A, B) = \exp(A)M\exp(B)$, sendo esta definida em um aberto de $\mathcal{A}(n) \times \mathcal{A}(n)$. O espaço tangente no ponto M é $T(O_\Sigma^+)_M = \{AM + MB : A, B \in \mathcal{A}(n)\}$.*

Prova: Primeiro precisamos verificar que $\psi_M(\mathcal{A}(n) \times \mathcal{A}(n)) \subset O_\Sigma^+$. Dadas matrizes anti-simétricas A e B, é claro que elas comutam com as suas respectivas transpostas, e pelo Lema 21 do Apêndice, temos que $\exp(A)$ e $\exp(B)$ estão em $\mathcal{SO}(n)$. Escrevendo $M = U\Sigma V$, segue que

$$\psi_M(A, B) = \exp(A)M\exp(B) = (\exp(A)U)\Sigma(V\exp(B)) \in O_\Sigma^+.$$

É claro que ψ_M é suave, $\psi_M(0,0) = M$ e $\dim(\mathcal{A}(n) \times \mathcal{A}(n)) = n^2 - n = \dim(T(O_\Sigma^+)_M)$. Pelo Lema 18 no Apêndice, para provarmos que ψ_M é uma carta, basta mostrarmos que

$$d(\psi_M)_{(0,0)} : \mathcal{A}(n) \times \mathcal{A}(n) \to T(O_\Sigma^+)_M$$

é um isomorfismo. Como $\mathcal{A}(n) \times \mathcal{A}(n)$ e $T(O_\Sigma^+)_M$ tem a mesma dimensão, isso é equivalente a mostrar que $d(\psi_M)_{(0,0)}$ é injetiva. Temos que

$$\begin{aligned} d(\psi_M)_{(0,0)}(A,B) &= \frac{d}{dt}\psi_M(tA,tB)|_{t=0} = \frac{d}{dt}\exp(tA)M\exp(tB)|_{t=0} \\ &= \{\exp(tA)AM\exp(tB) + \exp(tA)M\exp(tB)B\}|_{t=0} \\ &= AM + MB. \end{aligned}$$

Supondo $d(\psi_M)_{(0,0)}(A,B) = 0$, temos que

$$\begin{aligned} d(\psi_M)_{(0,0)}(A,B) = 0 &\Longrightarrow AM + MB = 0 \Longrightarrow AU\Sigma V + U\Sigma VB = 0 \\ &\Longrightarrow U^T AU\Sigma + \Sigma VBV^T = 0. \end{aligned}$$

Fazendo $C = U^T AU$ e $D = VBV^T$, obtemos

$$d(\psi_M)_{(0,0)}(A,B) = 0 \Longrightarrow \Sigma D + C\Sigma = 0 \Longrightarrow \sigma_i d_{ij} + c_{ij}\sigma_j = 0.$$

Como C e D são anti-simétricas, trocando os índices i e j, segue que $\sigma_j d_{ij} + c_{ij}\sigma_i = 0$. O sistema de equações

$$\sigma_i d_{ij} + \sigma_j c_{ij} = 0$$

$$\sigma_j d_{ij} + \sigma_i c_{ij} = 0,$$

para $i \neq j$, tem determinante nas variáveis c_{ij}, d_{ij} igual a $\sigma_i^2 - \sigma_j^2 \neq 0$. Segue que para $i \neq j$ temos $c_{ij} = d_{ij} = 0$, portanto $C = D = 0$, logo $A = B = 0$. Concluímos que $d(f_M)_{(0,0)}$ é injetiva.

Como ψ_M é uma carta em torno de M com $d(\psi_M)_{(0,0)}(A,B) = AM + MB$, é claro que $T(O_\Sigma^+)_M = \{AM + MB : A, B \in \mathcal{A}(n)\}$. ■

A prova destes lemas completa a demonstração do teorema.

Capítulo 3

A inclusão $F(O_\Sigma^+) \subset \mathcal{P}$

Lembramos que

$$\begin{aligned} F :& O_\Sigma^+ \to \mathbb{R}^n \\ & M \mapsto \operatorname{diag}(M), \end{aligned}$$

e $\mathcal{P}$ é o politopo convexo em $\mathbb{R}^n$ cujos vértices são pontos da forma $\{E\pi\sigma\}$, onde π é uma permutação e E é uma matriz ortogonal diagonal com determinante 1.

Dada uma matriz diagonal D com $d_i \in \{1, 0, -1\}$ para $1 \leq i \leq n$, onde $d = \operatorname{diag}(D)$, definimos a função

$$\begin{aligned} F_D :& O_\Sigma^+ \to \mathbb{R} \\ & M \mapsto F(M) \cdot d = \operatorname{tr}(MD). \end{aligned}$$

Neste capítulo mostraremos que $F(O_\Sigma^+) \subset \mathcal{P}$. Primeiro caracterizaremos os pontos críticos de F e veremos como estes estão relacionados com os pontos críticos de F_D. Depois exibiremos propriedades dos pontos críticos de F que permitirão calcular o valor da função F_D nesses pontos. Usando isso, provaremos que $F(O_\Sigma^+) \subset \mathcal{P}$.

O lema abaixo caracteriza os pontos críticos de F.

Lema 6 *Um ponto $M \in O_\Sigma^+$ é ponto crítico de F se e somente se existe uma matriz diagonal não nula D com $d_i \in \{1, 0, -1\}$ para $1 \leq i \leq n$ tal que MD e DM são simétricas, onde $d = \operatorname{diag}(D)$.*

Prova: Lembramos que $T(O_\Sigma^+)_M = \{AM + MB : A, B \in \mathcal{A}(n)\}$ e $dF_M : T(O_\Sigma^+)_M \to \mathbb{R}^n$. Se M é ponto crítico de F então dF_M não é sobrejetiva. Denotando por $W = dF_M(T(O_\Sigma^+)_M)$, temos que $\dim(W) < n$, logo existe um $d' \in W^\perp$ não nulo. Como F_D é a restrição de uma função linear definida num aberto contendo O_Σ^+, temos que $dF_M = F$ para qualquer $M \in O_\Sigma^+$, portanto $d' \cdot dF_M(AM + MB) = d' \cdot F(AM + MB) = 0$ para quaisquer matrizes anti-simétricas A e B. Definindo uma matriz diagonal D' com $\operatorname{diag}(D') = d'$, temos que $d' \cdot F(AM + MB) = \operatorname{tr} D'(AM + MB) = 0$ para quaisquer matrizes anti-simétricas A e B. Fazendo $A = 0$ e depois $B = 0$, obtemos $\mathrm{tr}(D'MB) = \mathrm{tr}(D'M)B = 0$ e $\mathrm{tr}(D'AM) = \mathrm{tr}(MD')A = 0$ para quaisquer matrizes ant- simétricas A e B. Lembrando que $\langle M, N\rangle = \mathrm{tr}(MN^T)$ é um produto interno em $\mathcal{M}(n)$ tal que $(\mathcal{S}(n))^\perp = \mathcal{A}(n)$ e $(\mathcal{A}(n))^\perp = \mathcal{S}(n)$, concluímos que MD' e $D'M$ são simétricas.

Como MD' e $D'M$ são simétricas, valem as equações $MD' = D'M^T$ e $D'M = M^TD'$. Olhando a entrada ij, obtemos $m_{ij}d'_j = d'_i m_{ji}$ e $m_{ij}d'_i = d'_j m_{ji}$. Multiplicando estas equações conseguimos $(d'_i)^2 m_{ij}m_{ji} = (d'_j)^2 m_{ij}m_{ji}$. Assim, $(d'_i)^2 \neq (d'_j)^2$ implica $m_{ij}m_{ji} = 0$. Usando as primeiras equações acima concluímos que $(d'_i)^2 \neq (d'_j)^2$ implica $m_{ij} = m_{ji} = 0$.

Denotemos por c a primeira entrada diagonal não nula de D'. Isso é possível pois $\mathrm{diag}(D') = d' \neq 0$. Definimos uma matriz diagonal D com $d = \mathrm{diag}(D)$ da seguinte forma

$$\begin{array}{lll} d_i = 0, & \text{se} & (d'_i)^2 \neq c^2 \\ d_i = 1, & \text{se} & d'_i = c \\ d_i = -1, & \text{se} & d'_i = -c. \end{array}$$

Provaremos que DM é simétrica. Para ver que MD é simétrica, a prova é análoga. Temos que DM é simétrica se e somente se a equação $DM - M^TD = 0$ é satisfeita. A entrada ij de $DM - M^TD$ é $d_i m_{ij} - d_j m_{ji}$. Se $d_i = d_j = 0$ esta equação é satisfeita. Se $(d_i)^2 \neq (d_j)^2$ então $(d'_i)^2 \neq (d'_j)^2$, portanto $m_{ij} = m_{ji} = 0$ e a equação é satisfeita. Se $(d_i)^2 = (d_j)^2 = 1$, existem as possibilidades $d_i = d_j$ e $d_i = -d_j$. Nestes casos, é fácil ver que as igualdades $m_{ij}d'_j = d'_i m_{ji}$ e $m_{ij}d'_i = d'_j m_{ji}$ garantem que a equação $DM - M^TD = 0$ é satisfeita. Esses são todos os casos, portanto DM é simétrica.

Agora mostraremos a recíproca. Suponhamos que D é uma matriz diagonal não nula com $d_i \in \{1, 0, -1\}$ para $1 \leq i \leq n$ tal que MD e DM são simétricas, onde $d = \mathrm{diag}(D)$. Temos que

$$d \cdot dF_M(AM + MB) = \mathrm{tr}D(AM + MB) = \mathrm{tr}(MD)A + \mathrm{tr}(DM)B = 0,$$

pois DM e MD são simétricas e A e B anti-simétricas, portanto

$$\dim(dF_M(T(O_\Sigma^+)_M)) < n,$$

consequentemente dF_M não é sobrejetiva. Concluímos que M é ponto crítico de F. ■

O lema abaixo mostra que os pontos críticos de F e de F_D estão relacionados da seguinte forma: M é ponto crítico de F se e somente se M é ponto crítico de alguma F_D.

Lema 7 *Dada uma matriz diagonal D com $d_i \in \{1, 0, -1\}$ para $1 \leq i \leq n$, onde $d = \mathrm{diag}(D)$, temos que $M \in O_\Sigma^+$ é ponto crítico de F_D se e somente se MD e DM são simétricas.*

Prova: Basta notarmos que a derivada $d(F_D)_M : T(O_\Sigma^+)_M \to \mathbb{R}$ é dada por

$$\begin{aligned} d(F_D)_M(AM + MB) &= F_D(AM + MB) \\ &= F(AM + MB) \cdot d \\ &= \operatorname{tr}(AM + MB)D \\ &= \operatorname{tr}A(MD) + \operatorname{tr}(DM)B. \end{aligned}$$

Utilizando um argumento análogo ao do lema anterior, concluímos que $d(F_D)_M = 0$ se e somente se MD e DM são simétricas. ■

O lema seguinte será usado logo no que se segue.

Lema 8 *Dadas uma matriz M e uma matriz diagonal não nula D tal que MD e DM são simétricas, o espaço vetorial gerado por $\{e_k : d_k^2 = c^2\}$ é invariante por M e M^T, onde $d = \operatorname{diag}(M)$ e c é uma entrada diagonal não nula de D.*

Prova: Como DM e MD são simétricas, temos que $MD = DM^T$ e $DM = M^T D$, portanto $MD^2 = (MD)D = D(M^T D) = D^2 M$. Analogamente $M^T D^2 = D^2 M^T$. Denotando por V o espaço vetorial gerado por $\{e_k : d_k^2 = c^2\}$, é fácil ver que $V = \{v : D^2 v = c^2 v\}$. Se $v \in V$, então $D^2 M v = MD^2 v = c^2 M v$, portanto $Mv \in V$, segue que V é invariante por M. Analogamente, usando que $M^T D^2 = D^2 M^T$, segue que V é invariante por M^T. ■

Os dois lemas seguintes são muito importantes. Através deles será possível calcular o valor de F_D nos seus pontos críticos.

Lema 9 *Dado um ponto crítico $M = U\Sigma V \in O_\Sigma^+$ de F_D, temos que $D' = VDU$ é diagonal com $d'_i \in \{1, 0, -1\}$ para $1 \le i \le n$, onde $d' = \operatorname{diag}(D')$. Além disso, o número de coordenadas 0 de d é igual ao número de coordenadas 0 de d', onde $d = \operatorname{diag}(D)$.*

Prova: Como M é ponto crítico de F_D, do Lema 7, segue que MD e DM são simétricas. Como DM é simétrica, temos que

$$\begin{aligned} M^T D = DM \Longrightarrow V^T \Sigma U^T D = DU\Sigma V \Longrightarrow \Sigma(VDU)^T = (VDU)\Sigma \\ \Longrightarrow \Sigma D'^T = D'\Sigma \Longrightarrow (D'\Sigma)^T = D'\Sigma, \end{aligned}$$

portanto $D'\Sigma$ é simétrica. Analogamente, usando que MD é simétrica, segue que $\Sigma D'$ é simétrica. Como $D'\Sigma$ e $\Sigma D'$ são simétricas e Σ é diagonal, pelo Lema 8, temos que para cada σ_i o espaço gerado pelo conjunto $\{e_k : \sigma_k^2 = \sigma_i^2\}$ é invariante por D'. Como $\sigma_i \ne \sigma_j$ para $i \ne j$, concluímos que D' é diagonal. Notando que $(D')^2$ e D^2 são conjugadas, pois

$$(D')^2 = D'D'^T = VDUU^T DV^T = VD^2V^T,$$

sabemos que $(D')^2$ e D^2 tem os mesmos autovalores, portanto $(D')^2$ só possui entradas 1 e 0 na diagonal e tem o mesmo número de 0 na diagonal que D^2. Logo as entradas da diagonal de D' são $1, 0$ e -1, e tem o mesmo número de 0 na diagonal que D. ■

Lema 10 *Dado um ponto crítico* $M = U\Sigma V \in O^+_\Sigma$ *de* F_D *tal que todas as coordenadas de* d *são não nulas, temos que o número de coordenadas* -1 *de* d *e o número de coordenadas* -1 *de* d' *tem a mesma paridade, onde* $D' = VDU$, $d = \mathrm{diag}(D)$ *e* $d' = \mathrm{diag}(D')$.

Prova: Como U e V são ortogonais, segue que $\det(U)\det(V) = \pm 1$. Como M tem determinante positivo, temos que $\det(U)\det(V) = 1$, portanto $\det(D') = \det(VDU) = \det(U)\det(D)\det(V) = \det D$. Pelo lema acima, sabemos D' é diagonal com $d'_i \in \{1, -1\}$ para $1 \leq i \leq n$. Denotando por a o número de coordenadas -1 de d e por b o número de coordenadas -1 de d', temos que $(-1)^a = \det(D) = \det(D') = (-1)^b$, e isso implica a conclusão. ■

Agora iremos provar a primeira parte do Teorema de Thompson usando os lemas acima.

Primeiro daremos uma descrição do politopo $\mathcal{P}$ através das suas faces. Uma prova de que a descrição de $\mathcal{P}$ pelos vértices coincide com a descrição pelas faces pode ser encontrado em [2].

Definimos

$$\begin{aligned}
u_1 &= e_1 \\
&\vdots \\
u_k &= e_1 + \ldots + e_k \\
&\vdots \\
u_{n-2} &= e_1 + \ldots + e_{n-2} \\
u_{n-1} &= e_1 + \ldots + e_{n-1} - e_n \\
u_n &= e_1 + \ldots + e_{n-1} + e_n.
\end{aligned}$$

Para n par, temos que $\mathcal{P}$ é o conjunto dos $x \in \mathbb{R}^n$ tais que para toda permutação π, toda matriz ortogonal diagonal E com determinante 1 e todo $1 \leq k \leq n$ vale

$$|x \cdot E\pi(u_k)| \leq \sigma \cdot u_k.$$

Para n ímpar, temos que $\mathcal{P}$ é o conjunto dos $x \in \mathbb{R}^n$ tais que para toda permutação π, toda matriz ortogonal diagonal E com determinante 1 e todo $1 \leq k \leq n-2$ vale

$$\begin{aligned}
&|x \cdot E\pi(u_k)| \leq \sigma \cdot u_k \\
&\sigma \cdot u_{n-1} \leq x \cdot E\pi(u_k) \leq \sigma \cdot u_n \\
&\sigma \cdot u_n \leq x \cdot E\pi(u_k) \leq \sigma \cdot u_{n-1}.
\end{aligned}$$

Podemos ver $\mathcal{P}$ como o conjunto dos pontos $x \in \mathbb{R}^n$ que são soluções das seguintes inequações:

- para todo vetor $d = (d_1, \ldots, d_n) \in \mathbb{R}^n$ com $d_i \in \{1, 0, -1\}$ para $1 \leq i \leq n$, com $k \leq n$ coordenadas não nulas, vale

$$-(\sigma_1 + \ldots + \sigma_k) \leq x \cdot d \leq \sigma_1 + \ldots + \sigma_k$$

- para todo vetor $d = (d_1, \ldots, d_n) \in \mathbb{R}^n$ com $d_i \in \{1, -1\}$ para $1 \leq i \leq n$, tendo um número par de coordenadas -1, vale

$$-(\sigma_1 + \ldots + \sigma_n) \leq x \cdot d \leq \sigma_1 + \ldots + \sigma_n, \text{ se } n \text{ é par}$$

$$-(\sigma_1 + \ldots + \sigma_{n-1} - \sigma_n) \leq x \cdot d \leq \sigma_1 + \ldots + \sigma_n, \text{ se } n \text{ é ímpar}$$

- para todo vetor $d = (d_1, \ldots, d_n) \in \mathbb{R}^n$ com $d_i \in \{1, -1\}$ para $1 \leq i \leq n$, tendo um número ímpar de coordenadas -1, vale

$$-(\sigma_1 + \ldots + \sigma_{n-1} - \sigma_n) \leq x \cdot d \leq \sigma_1 + \ldots + \sigma_{n-1} - \sigma_n, \text{ se } n \text{ é par}$$

$$-(\sigma_1 + \ldots + \sigma_n) \leq x \cdot d \leq \sigma_1 + \ldots + \sigma_{n-1} - \sigma_n, \text{ se } n \text{ é ímpar}$$

Usando essa descrição do politopo $\mathcal{P}$ seremos capazes de provar o teorema abaixo.

Teorema 11 *$F(O_\Sigma^+) \subset \mathcal{P}$.*

Prova: Para mostrarmos que $F(O_\Sigma^+) \subset \mathcal{P}$, é suficiente verificarmos que os pontos de $F(O_\Sigma^+)$ são soluções das inequações que descrevem as faces de $\mathcal{P}$.

Dado um vetor $d \in \mathbb{R}^n$ com $d_i \in \{1, 0, -1\}$ para $1 \leq i \leq n$, denotando por D a matriz diagonal com $\operatorname{diag}(D) = d$, como $F_D : O_\Sigma^+ \to \mathbb{R}$ é contínua e O_Σ^+ é compacto, existem $M^+ = U^+ \Sigma V^+$ e $M^- = U^- \Sigma V^-$ em O_Σ^+ tais que

$$\max_{O_\Sigma^+} F_D = F_D(M^+), \quad \text{e} \quad \min_{O_\Sigma^+} F_D = F_D(M^-).$$

Agora, o valor de $F_D(M^+)$ é

$$F(M^+) \cdot d = \operatorname{tr}(M^+ D) = \operatorname{tr}(U^+ \Sigma V^+ D) = \operatorname{tr}(V^+ D U^+)\Sigma = \operatorname{tr}((D')^+ \Sigma),$$

onde $(D')^+ = V^+ D U^+$. Analogamente obtemos que $F_D(M^-) = \operatorname{tr}((D')^- \Sigma)$, onde $(D')^- = V^- D U^-$. Denotando por $(d')^+ = \operatorname{diag}((D')^+)$ e $(d')^- = \operatorname{diag}((D')^-)$, como M^+ e M^- são pontos críticos de F_D, pelo Lema 9, segue que $(D')^+$ e $(D')^-$ são diagonais com $(d')_i^+, (d')_i^- \in \{1, 0, -1\}$ para $1 \leq i \leq n$, e além disso o número de coordenadas não nulas de $(d')^+, (d')^-$ e d são iguais a, digamos, $k \leq n$. Logo existem índices $i_1^+, \ldots, i_k^+, i_1^-, \ldots, i_k^- \in \mathbb{N}$ tais que

$$F_D(M^+) = \pm\sigma_{i_1^+} + \ldots + \pm\sigma_{i_k^+} \text{ e } F_D(M^-) = \pm\sigma_{i_1^-} + \ldots + \pm\sigma_{i_k^-}.$$

Portanto $F_D(M^+) \leq \sigma_1 + \ldots + \sigma_k$ e $F_D(M^-) \geq -(\sigma_1 + \ldots + \sigma_k)$.

Agora suponhamos que todas as coordenadas de d são não nulas e o número de coordenadas de d iguais a -1 é ímpar. Pelo Lema 10, segue que o número de coordenadas -1 de $(d')^+$ também é ímpar.

Lembrando que $\sigma_1 > \ldots > \sigma_n$, o máximo de uma soma da forma $\pm\sigma_1 + \ldots + \pm\sigma_n$ com um número ímpar de sinais negativos é quando for igual a $\sigma_1 + \ldots + \sigma_{n-1} - \sigma_n$, portanto $F_D(M^+) \leq \sigma_1 + \ldots + \sigma_{n-1} - \sigma_n$.

O mínimo que uma soma da forma $\pm\sigma_1 + \ldots + \pm\sigma_n$ com um número ímpar de sinais negativos é quando for igual a

$$
\begin{aligned}
&-(\sigma_1 + \ldots + \sigma_{n-1} - \sigma_n), \text{ se } n \text{ é par} \\
&-(\sigma_1 + \ldots + \sigma_n), \text{ se } n \text{ é ímpar.}
\end{aligned}
$$

Então

$$
\begin{aligned}
&F_D(M^-) \geq -(\sigma_1 + \ldots + \sigma_{n-1} - \sigma_n), \text{ se } n \text{ é par} \\
&F_D(M^-) \geq -(\sigma_1 + \ldots + \sigma_n), \text{ se } n \text{ é ímpar.}
\end{aligned}
$$

Se o número de coordenadas -1 de d é par, seguindo os mesmos argumentos acima, segue que

$$
\begin{aligned}
&F_D(M^+) \leq \sigma_1 + \ldots + \sigma_{n-1} - \sigma_n \\
&F_D(M^-) \geq -(\sigma_1 + \ldots + \sigma_n), \text{ se } n \text{ é par} \\
&F_D(M^-) \geq -(\sigma_1 + \ldots + \sigma_{n-1} - \sigma_n), \text{ se } n \text{ é ímpar.}
\end{aligned}
$$

Para todo $M \in O_\Sigma^+$, temos que $\min_{O_\Sigma^+} F_D = F_D(M^-) \leq F_D(M) \leq F_D(M^+) = \max_{O_\Sigma^+} F_D$. Segue o resultado. ∎

Capítulo 4

A inclusão $\mathcal{P} \subset F(O_\Sigma^+)$

Dados vetores $d = (d_1, \ldots, d_n), d' = (d'_1, \ldots, d'_n) \in \mathbb{R}^n$ tais que $d_i, d'_i \in \{1, 0, -1\}$ para $1 \leq i \leq n$, chamamos o conjunto

$$H_{d,d'} = \{x \in \mathbb{R}^n : x \cdot d = \sigma \cdot d'\}$$

de *hiperplano crítico*. Se d' é um vetor da forma $\pm(1, \ldots, 1, 0, \ldots, 0), \pm(1, \ldots, 1)$ ou $\pm(1, \ldots, 1, -1)$, dizemos que $H_{d,d'}$ é um *hiperplano crítico extremo.*

Neste capítulo mostraremos que $\mathcal{P} = F(O_\Sigma^+)$. Primeiro veremos que a fronteira de $F(O_\Sigma^+)$ está contida na união dos hiperplanos críticos. Depois faremos um estudo da segunda derivada da função F_D, e através disso provaremos que os pontos da fronteira de $F(O_\Sigma^+)$ que estão num único hiperplano crítico não extremo estão necessariamente no interior de $F(O_\Sigma^+)$. Por fim construiremos um conjunto conexo e denso em $\mathcal{P}$ que não possui pontos em comum com a fronteira de $F(O_\Sigma^+)$, a saber os pontos do interior do politopo $\mathcal{P}$ menos os pontos que estão em dois ou mais hiperplanos críticos, e usando um argumento de conexidade concluiremos que $\mathcal{P} = F(O_\Sigma^+)$.

Lema 12 *A fronteira de $F(O_\Sigma^+)$ está contida na união de todos os hiperplanos críticos.*

Prova: Primeiro mostraremos que a fronteira de $F(O_\Sigma^+)$ está contida na imagem dos pontos críticos de F. Se M é um ponto regular de F, do Lema 20 no Apêndice, segue que $F(M)$ está no interior de $F(O_\Sigma^+)$, portanto todo ponto na fronteira de $F(O_\Sigma^+)$ é necessariamente imagem de um ponto crítico de F. Nessa última frase usamos o fato de que $F(O_\Sigma^+)$ é fechado.

Agora veremos que a imagem dos pontos críticos de F está contida na união dos hiperplanos críticos. Dado um ponto crítico $M = U\Sigma V \in O_\Sigma^+$ de F, pelo Lema 7, existe uma matriz diagonal D com $d_i \in \{1, 0, -1\}$ para $1 \leq i \leq n$ tal que M é ponto crítico de F_D, onde $d = \operatorname{diag}(D)$. Temos que

$$F_D(M) = \operatorname{tr}(U\Sigma VD) = \operatorname{tr}(\Sigma(VDU)) = \operatorname{tr}(\Sigma D'),$$

onde $D' = VDU$. Pelo Lema 9, temos que D' é diagonal com $d'_i \in \{1, 0, -1\}$ para $1 \leq i \leq n$, onde $d' = \operatorname{diag}(D')$. Agora a equação $F_D(M) = \operatorname{tr}(\Sigma D')$ é equivalente a $F(M) \cdot d = \sigma \cdot d'$, portanto $H_{d,d'}$ é um hiperplano crítico e contém $F(M)$. ■

O lema seguinte é essencial para provarmos o próximo resultado. Faremos uma análise da segunda derivada da função F_D em seus pontos críticos que estão em hiperplanos não extremos.

Lema 13 *Se $M = U\Sigma V \in O_\Sigma^+$ é um ponto crítico de F_D e o hiperplano crítico $H_{d,d'}$ não é extremo, onde $D' = VDU$, $d = \operatorname{diag}(D)$ e $d' = \operatorname{diag}(D')$, existem matrizes $M^+, M^- \in O_\Sigma^+$ arbitrariamente próximas de M tais que*

$$F_D(M^+) > F_D(M) > F_D(M^-).$$

Prova: Dadas duas matrizes anti-simétricas A e B, pelo Teorema 1, para t definido num intervalo aberto contendo 0 suficientemente pequeno, a imagem da função $\psi_{A,B}(t) = \exp(tA)M\exp(tB)$ está contida em O_Σ^+. Temos que $F_D \circ \psi_{A,B}(t) = \operatorname{tr}[D\exp(tA)M\exp(tB)]$, portanto

$$\begin{aligned}
\frac{d}{dt}(F_D \circ \psi_{A,B})(t) &= \operatorname{tr}\left[D(\exp(tA)AM\exp(tB) + \exp(tA)M\exp(tB)B)\right] \\
\frac{d^2}{dt^2}(F_D \circ \psi_{A,B})(t) &= \operatorname{tr}\big[D(\exp(tA)A^2M\exp(tB) + \\
&\qquad +2\exp(tA)AM\exp(tB)B + \exp(tA)M\exp(tB)B^2)\big] \\
\frac{d^2}{dt^2}(F_D \circ \psi_{A,B})|_{t=0} &= \operatorname{tr}\left[D(A^2M + 2AMB + MB^2)\right].
\end{aligned}$$

Pela fórmula de Taylor para $F_D \circ \psi_{A,B}$ no ponto $t = 0$, temos que

$$F_D \circ \psi_{A,B}(t) = F_D(M) + t\frac{d}{dt}(F_D \circ \psi_{A,B})|_{t=0} + \frac{t^2}{2}\frac{d^2}{dt^2}(F_D \circ \psi_{A,B})|_{t=0} + t^2 g(t),$$

onde $\lim_{t\to 0} g(t) = 0$. Como M é ponto crítico de F_D, temos que $F_D(\frac{d}{dt}\psi_{A,B}|_{t=0}) = 0$, portanto

$$F_D \circ \psi_{A,B}(t) - F_D(M) = \frac{t^2}{2}\left[F_D(\frac{d^2}{dt^2}\psi_{A,B}|_{t=0}) + 2g(t)\right],$$

onde $\lim_{t\to 0} g(t) = 0$. Assim, para t suficientemente próximo de 0, o sinal de $F_D \circ \psi_{A,B}(t) - F_D(M)$ só depende do sinal de $F_D(\frac{d^2}{dt^2}\psi_{A,B}|_{t=0})$.

Encontraremos matrizes anti-simétricas A^+, A^-, B^+ e B^- tais que

$$F_D\left(\frac{d^2}{dt^2}\psi_{A^+,B^+}|_{t=0}\right) > 0 \qquad \text{e} \qquad F_D\left(\frac{d^2}{dt^2}\psi_{A^-,B^-}|_{t=0}\right) < 0.$$

Dessa forma, para t suficientemente próximo de 0, teremos

$$F_D(\psi_{A^+,B^+}(t)) - F_D(M) > 0, \quad \text{e} \quad F_D(\psi_{A^-,B^-}(t)) - F_D(M) < 0,$$

finalizando a prova do lema.

Temos que

$$\begin{aligned}
F_D\left(\frac{d^2}{dt^2}\psi_{A,B}|_{t=0}\right) &= \operatorname{tr}\left[D(A^2M + 2AMB + MB^2)\right] \\
&= \operatorname{tr}\left(DA^2U\Sigma V + 2DAU\Sigma VB + DU\Sigma VB^2\right) \\
&= \operatorname{tr}\left((VDU)U^TA^2U\Sigma + 2(VDU)U^TAU\Sigma VBV^T\right. \\
&\quad \left. +(VDU)\Sigma VB^2V^T\right) \\
&= \operatorname{tr}\left(D'(U^TA^2U)\Sigma + 2D'(U^TAU)\Sigma(VBV^T)\right. \\
&\quad \left. +D'\Sigma(VB^2V^T)\right).
\end{aligned}$$

Definindo $\tilde{A} = U^TAU$ e $\tilde{B} = VBV^T$, obtemos $\tilde{A}^2 = U^TA^2U$ e $\tilde{B}^2 = VB^2V^T$, portanto

$$F_D\left(\frac{d^2}{dt^2}\psi_{A,B}|_{t=0}\right) = \operatorname{tr}(D'\tilde{A}^2\Sigma) + 2\operatorname{tr}(D'\tilde{A}\Sigma\tilde{B}) + \operatorname{tr}(D'\Sigma\tilde{B}^2).$$

Como $A = U\tilde{A}U^T$ e $B = V^T\tilde{B}V$, para provar o resultado, basta encontrarmos matrizes anti-simétricas $\tilde{A}^+, \tilde{A}^-, \tilde{B}^+$ e $\tilde{B}^-$ tais que

$$\begin{aligned}
F_D\left(\frac{d^2}{dt^2}\psi_{\tilde{A}^+,\tilde{B}^+}|_{t=0}\right) &= \operatorname{tr}(D'(\tilde{A}^+)^2\Sigma) + 2\operatorname{tr}(D'\tilde{A}^+\Sigma\tilde{B}^+) + \operatorname{tr}(D'\Sigma(\tilde{B}^+)^2) > 0 \\
F_D\left(\frac{d^2}{dt^2}\psi_{\tilde{A}^-,\tilde{B}^-}|_{t=0}\right) &= \operatorname{tr}(D'(\tilde{A}^-)^2\Sigma) + 2\operatorname{tr}(D'\tilde{A}^-\Sigma\tilde{B}^-) + \operatorname{tr}(D'\Sigma(\tilde{B}^-)^2) < 0.
\end{aligned}$$

Fixados $i \neq j$ definimos a matriz $A = A^{ij}$ tendo os valores $A_{ij} = 1$, $A_{ji} = -1$, e 0 nas entradas restantes. Agora, escrevendo os vetores da base canônica de $\mathbb{R}^n$ como matrizes $n \times 1$, é fácil ver que $A = e_ie_j^T - e_je_i^T$, e portanto $A^2 = -e_ie_i^T - e_je_j^T$, logo $(A^2)_{ii} = (A^2)_{jj} = -1$ e é 0 nas outras entradas. Portanto

$$\operatorname{tr}(D'\Sigma A^2) = \operatorname{tr}(D'A^2\Sigma) = -d'_i\sigma_i - d'_j\sigma_j.$$

Além disso, também temos que

$$\begin{aligned}
A\Sigma A &= (e_ie_j^T - e_je_i^T)\Sigma(e_ie_j^T - e_je_i^T) \\
&= (\sigma_je_ie_j^T - \sigma_ie_je_i^T)(e_ie_j^T - e_je_i^T) \\
&= -\sigma_je_ie_i^T - \sigma_ie_je_j^T,
\end{aligned}$$

portanto concluímos que $(A\Sigma A)_{ii} = -\sigma_j$, $(A\Sigma A)_{jj} = -\sigma_i$ e é 0 nas outras entradas. Segue que $\mathrm{tr}(D'A\Sigma A) = -d_i'\sigma_j - d_j'\sigma_i$. Consequentemente

$$\begin{aligned}
F_D\left(\frac{d^2}{dt^2}\psi_{A^{ij},A^{ij}}|_{t=0}\right) &= \mathrm{tr}(D'A^2\Sigma) + 2\mathrm{tr}(D'A\Sigma A) + \mathrm{tr}(D'\Sigma A^2) \\
&= 2(-d_i'\sigma_i - d_j'\sigma_j) + 2(-d_i'\sigma_j - d_j'\sigma_i) \\
&= -2(\sigma_i(d_i' + d_j') + \sigma_j(d_i' + d_j')) = -2(\sigma_i + \sigma_j)(d_i' + d_j') \\
F_D\left(\frac{d^2}{dt^2}\psi_{A^{ij},-A^{ij}}|_{t=0}\right) &= \mathrm{tr}(D'A^2\Sigma) - 2\mathrm{tr}(D'A\Sigma A) + \mathrm{tr}(D'\Sigma A^2) \\
&= 2(-d_i'\sigma_i - d_j'\sigma_j) - 2(-d_i'\sigma_j - d_j'\sigma_i) \\
&= 2(\sigma_i(-d_i' + d_j') + \sigma_j(d_i' - d_j')) = 2(\sigma_i - \sigma_j)(d_j' - d_i').
\end{aligned}$$

Para terminarmos a demonstração, basta determinarmos índices i e j que troquem o sinal de $F_D\left(\frac{d^2}{dt^2}\psi_{A^{ij},A^{ij}}|_{t=0}\right)$ e $F_D\left(\frac{d^2}{dt^2}\psi_{A^{ij},-A^{ij}}|_{t=0}\right)$ em função do vetor d'.

Primeiro suponhamos que todas as entradas de d' são não nulas. Como $H_{d,d'}$ não é extremo, temos que d' é diferente dos vetores $\pm(1,\ldots,1)$ e $\pm(1,\ldots,1,-1)$. Estudaremos agora os possíveis casos para d'.

Se d' é da forma $(\ldots,1,\ldots,-1,\ldots,1,\ldots)$, denotando por d_{i_1}', d_{i_2}' e d_{i_3}' as coordenadas $1, -1$ e 1 em ordem, temos que

$$\begin{aligned}
F_D\left(\frac{d^2}{dt^2}\psi_{A^{i_1,i_2},-A^{i_1,i_2}}|_{t=0}\right) &= 2(d_{i_1}' - d_{i_2}')(\sigma_{i_2} - \sigma_{i_1}) = 4(\sigma_{i_2} - \sigma_{i_1}) < 0 \\
F_D\left(\frac{d^2}{dt^2}\psi_{A^{i_2,i_3},-A^{i_2,i_3}}|_{t=0}\right) &= 2(d_{i_2}' - d_{i_3}')(\sigma_{i_3} - \sigma_{i_2}) = -4(\sigma_{i_3} - \sigma_{i_2}) > 0.
\end{aligned}$$

Se d' é da forma $(\ldots,-1,\ldots,1,\ldots,-1,\ldots)$, é análogo.

Se d' é da forma $(\ldots,1,\ldots,-1,\ldots,-1)$, denotando por d_{i_1}', d_{i_2}' e d_{i_3}' as coordenadas $1, -1$ e -1 em ordem, temos que

$$\begin{aligned}
F_D\left(\frac{d^2}{dt^2}\psi_{A^{i_1,i_2},-A^{i_1,i_2}}|_{t=0}\right) &= 2(d_{i_1}' - d_{i_2}')(\sigma_{i_2} - \sigma_{i_1}) = 4(\sigma_{i_2} - \sigma_{i_1}) < 0 \\
F_D\left(\frac{d^2}{dt^2}\psi_{A^{i_2,i_3},A^{i_2,i_3}}|_{t=0}\right) &= -2(d_{i_2}' + d_{i_3}')(\sigma_{i_2} + \sigma_{i_3}) = 4(\sigma_{i_2} + \sigma_{i_3}) > 0.
\end{aligned}$$

Se d' é da forma $(\ldots,-1,\ldots,1,\ldots,1)$, é análogo.

Como d' é diferente dos vetores $\pm(1,\ldots,1)$ e $\pm(1,\ldots,1,-1)$, ele é de uma das formas acima. Isso resolve o problema no caso onde d' tem todas as coordenadas não nulas.

Agora suponhamos que d' tem alguma coordenada nula. Como $H_{d,d'}$ não é extremo, temos que d' é diferente dos vetores $\pm(1,\ldots,1,0,\ldots,0)$, portanto em algum momento aparece uma sequência 0 ± 1, ou melhor, d' é da forma

$(\ldots, 0, \pm 1, \ldots)$. Suponhamos, sem perda de generalidade, que d' é da forma $(\ldots, 0, 1, \ldots)$. Denotando por d'_{i_1}, d'_{i_2} as coordenadas 0 e 1 acima, temos que

$$\begin{aligned} F_D\left(\frac{d^2}{dt^2}\psi_{A^{i_1,i_2},-A^{i_1,i_2}}|_{t=0}\right) &= 2(d'_{i_1} - d'_{i_2})(\sigma_{i_2} - \sigma_{i_1}) = -2(\sigma_{i_2} - \sigma_{i_1}) > 0 \\ F_D\left(\frac{d^2}{dt^2}\psi_{A^{i_2,i_3},A^{i_2,i_3}}|_{t=0}\right) &= -2(d'_{i_1} + d'_{i_2})(\sigma_{i_1} + \sigma_{i_2}) = -2(\sigma_{i_1} + \sigma_{i_2}) < 0. \end{aligned}$$

Isso finaliza a prova do lema. ■

Agora veremos que os hiperplanos críticos não extremos não contém pontos da fronteira de $F(O_\Sigma^+)$ que estão num único hiperplano crítico.

Lema 14 *Se $M \in O_\Sigma^+$ é um ponto crítico de F tal que $F(M)$ está num único hiperplano crítico que não é extremo então $F(M) \in \operatorname{int}(F(O_\Sigma^+))$.*

Prova: Denotemos por $H_{d,d'}$ o único hiperplano crítico que contém $F(M)$. Para cada hiperplano crítico $H_{e,e'}$ com $e \neq d$ e $e' \neq d'$, como $F(M) \notin H_{e,e'}$ e $H_{e,e'}$ é fechado, existe uma bola

$$B_{e,e'} = \{x \in \mathbb{R}^n : |x - F(M)| < r_{e,e'}\}$$

com centro em $F(M)$ tal que $B_{e,e'} \cap H_{e,e'} = \emptyset$. Temos que os hiperplanos críticos existem em número finito, pois os vetores c, c' que determinam um hiperplano crítico $H_{c,c'}$ só podem ter coordenadas $1, 0$ e -1. Repetindo esse processo para cada hiperplano crítico diferente de $H_{d,d'}$, obtemos uma bola

$$B = \{x \in \mathbb{R}^n : |x - F(M)| < r\}$$

com centro em $F(M)$ tal que $B \cap H_{e,e'} = \emptyset$ para cada hiperplano crítico $H_{e,e'}$ com $e \neq d$ e $e' \neq d'$.

Para mostrarmos que $F(M) \in \operatorname{int}(F(O_\Sigma^+))$, é suficiente verificarmos que $B \subset F(O_\Sigma^+)$. Definindo

$$\begin{aligned} B^+ &= \{x \in B : x \cdot d > \sigma \cdot d'\}, \\ B^0 &= \{x \in B : x \cdot d = \sigma \cdot d'\}, \\ B^- &= \{x \in B : x \cdot d < \sigma \cdot d'\}, \end{aligned}$$

teremos que $B = B^+ \cup B^0 \cup B^-$, e além disso B^0 está no fecho de B^+ e B^-. Como $F(O_\Sigma^+)$ é fechado, para demonstrar que $B \subset F(O_\Sigma^+)$, basta provar que $B^+, B^- \subset F(O_\Sigma^+)$.

Vejamos que $B^+ \subset F(O_\Sigma^+)$. Para provar que $B^- \subset F(O_\Sigma^+)$ é análogo. A estratégia é provar que $B^+ \cap \operatorname{int}(F(O_\Sigma^+)) = B^+$, isso garante que $B^+ \subset \operatorname{int}(F(O_\Sigma^+)) \subset F(O_\Sigma^+)$. Podemos escrever

$$B^+ = \left[B^+ \cap \operatorname{int}(F(O_\Sigma^+))\right] \cup \left[B^+ \cap \operatorname{fr}(F(O_\Sigma^+))\right] \cup \left[B^+ \cap \operatorname{ext}(F(O_\Sigma^+))\right].$$

Se provarmos que $B^+ \cap \mathrm{fr}(F(O_\Sigma^+)) = \emptyset$, teremos que B^+ é união disjuntos de dois abertos disjuntos em B^+, a saber $B^+ \cap \mathrm{int}(F(O_\Sigma^+))$ e $B^+ \cap \mathrm{ext}(F(O_\Sigma^+))$. Como B^+ é conexo, teremos que um destes dois conjuntos é vazio. Se provarmos que $B^+ \cap \mathrm{int}(F(O_\Sigma^+)) \neq \emptyset$, poderemos concluir que $B^+ \cap \mathrm{int}(F(O_\Sigma^+)) = B^+$, portanto $B^+ \subset F(O_\Sigma^+)$.

Primeiro mostraremos que $B^+ \cap \mathrm{fr}(F(O_\Sigma^+)) = \emptyset$. Por construção, temos que B^+ não tem interseção com nenhum hiperplano crítico. Pelo Lema 12, temos que $\mathrm{fr}(F(O_\Sigma^+))$ está contida na união dos hiperplanos críticos, portanto B^+ não tem interseção com $\mathrm{fr}(F(O_\Sigma^+))$, ou seja, $B^+ \cap \mathrm{fr}(F(O_\Sigma^+)) = \emptyset$.

Agora provaremos que $B^+ \cap \mathrm{int}(F(O_\Sigma^+)) \neq \emptyset$. Escrevendo $M = U\Sigma V$, como é ponto crítico de F, seguindo o segundo parágrafo do Lema 12, temos que $F(M)$ está no hiperplano crítico $H_{d,d''}$, onde $D'' = VDU$ e $d'' = \mathrm{diag}(D'')$. Como $F(M)$ está num único hiperplano crítico, concluímos que $d'' = d' = \mathrm{diag}(VDU)$. Como $H_{d,d'}$ não é extremo, pelo Lema 13, existem matrizes M^+ arbitrariamente próximas de M tais que $F_D(M) < F_D(M^+)$. Esta inequação pode ser escrita na forma

$$F(M) \cdot d < F(M^+) \cdot d.$$

Como $F(M) \cdot d = \sigma \cdot d'$, concluímos que

$$F(M^+) \cdot d > \sigma \cdot d',$$

portanto $B^+ \cap F(O_\Sigma^+) \neq \emptyset$. Como $B^+ \cap \mathrm{fr}(F(O_\Sigma^+)) = \emptyset$, temos que $B^+ \cap \mathrm{int}(F(O_\Sigma^+)) \neq \emptyset$. ■

Usando os lemas acima, mostraremos o teorema principal do texto.

Teorema 15 $F(O_\Sigma^+) = \mathcal{P}$

Prova: A descrição de $\mathcal{P}$ pelas faces deixa claro que $\mathrm{int}(\mathcal{P})$ é conexo. Denotemos por A o conjunto formado pelos pontos de $\mathbb{R}^n$ que estão na interseção de dois ou mais hiperplanos críticos.

Vejamos que $\mathrm{int}(\mathcal{P}) - A$ é conexo. Não é difícil ver que a interseção de dois ou mais hiperplanos críticos forma um subespaço de $\mathbb{R}^n$ de codimensão maior ou igual a dois. Usando o fato de que ao retirarmos de um conjunto conexo em $\mathbb{R}^n$ os seus pontos que estão num subespaço de codimensão maior ou igual a dois o conjunto restante é conexo, e que os hiperplanos críticos existem em número finito, concluímos que o conjunto $\mathrm{int}(\mathcal{P}) - A$ é conexo.

Agora mostraremos que $\mathrm{int}(\mathcal{P}) - A \subset \mathrm{int} F(O_\Sigma^+)$. Como no lema anterior, usando a conexidade de $\mathrm{int}(\mathcal{P}) - A$, só precisamos verificar que

$$(\mathrm{int}(\mathcal{P}) - A) \cap \mathrm{fr}(F(O_\Sigma^+)) = \emptyset \quad \text{e} \quad (\mathrm{int}(\mathcal{P}) - A) \cap \mathrm{int}(F(O_\Sigma^+)) \neq \emptyset.$$

Vejamos que $(\mathrm{int}(\mathcal{P}) - A) \cap \mathrm{fr}(F(O_\Sigma^+)) = \emptyset$. Já mostramos no Lema 12 que $\mathrm{fr}(F(O_\Sigma^+))$ está contida na união dos hiperplanos críticos. Vimos no Lema 14 que os pontos de $F(O_\Sigma^+)$ que estão num único hiperplano crítico que não é extremo estão em $\text{int}(F(O_\Sigma^+))$. Portanto os pontos de $\text{fr}(F(O_\Sigma^+))$ estão nos hiperplanos críticos extremos ou são pontos que estão em dois ou mais hiperplanos críticos, ou seja, são pontos de $\text{fr}(\mathcal{P})$ ou de A. Segue que em $\text{int}(\mathcal{P}) - A$ não existem pontos de $\text{fr}(F(O_\Sigma^+))$.

Para provar $(\text{int}(\mathcal{P}) - A) \cap \text{int}(F(O_\Sigma^+)) \neq \emptyset$, basta escolhermos um ponto regular M da função F, pois nesse caso o Lema 20 do Apêndice garante que $F(M) \in \text{int}(F(O_\Sigma^+))$.

Finalmente provaremos que $F(O_\Sigma^+) = \mathcal{P}$. Se mostrarmos que $\text{int}(\mathcal{P}) - A$ é denso em $\text{int}(\mathcal{P})$, tomando o fecho em $\text{int}(\mathcal{P}) - A \subset \text{int}F(O_\Sigma^+)$, poderemos concluir que $\text{int}(\mathcal{P}) \subset F(O_\Sigma^+)$. Lembrando que $F(O_\Sigma^+)$ é fechado, e tomando o fecho mais uma vez, segue que $\mathcal{P} \subset F(O_\Sigma^+)$. Portanto, só nos resta mostrar que $\text{int}(\mathcal{P}) - A$ é denso em $\text{int}(\mathcal{P})$. Isso de fato acontece, pois A é união finita das interseções dois a dois, três a três, etc. de hiperplanos críticos, que são conjuntos com interior vazio em $\mathbb{R}^n$. ■

Apêndice

Valores Singulares

Lembramos o enunciado do Teorema espectral.

Teorema 16 *Se M é uma matriz real simétrica $n \times n$, então $M = U\Lambda U^T$, onde U é uma matriz real ortogonal $n \times n$ e Λ é uma matriz diagonal cujas entradas diagonais são os autovalores de M.*

Se M é uma matriz real $n \times n$, então MM^T e M^TM são matrizes simétricas nas quais podemos aplicar o Teorema espectral. Não é difícil verificar que os autovalores de MM^T e de M^TM são os mesmos e positivos. Se $\sigma_1^2, \dots, \sigma_n^2$ são os autovalores de MM^T, os números $\sigma_1, \dots, \sigma_n$ são chamados de valores singulares de M. Existe uma forma equivalente de definir valores singulares. Um número positivo σ é chamado valor singular de M se existem vetores u e v em $\mathbb{R}^n$ tais que $M^Tu = \sigma v$ e $Mv = \sigma u$. Os vetores u e v são chamados vetores singulares a esquerda e a direita respectivamente.

Agora enunciamos o Teorema da decomposição em valores singulares para matrizes quadradas. Existe também uma versão deste teorema para matrizes $m \times n$.

Teorema 17 *Se M é uma matriz real $n \times n$, então $M = U\Sigma V$, onde U e V são matrizes reais ortogonais $n \times n$ e Σ é uma matriz diagonal cujas entradas diagonais são os valores singulares de M.*

Prova: Denotando por S^{n-1} a esfera unitária em $\mathbb{R}^n$, considere a função f de $S^{n-1} \times S^{n-1}$ em $\mathbb{R}$ dada por $f(u, v) = Mu \cdot v$. Como esta função é contínua e definida num compacto ela atinge seu valor máximo num ponto (u_1, v_1). O número real $f(u_1, v_1)$ é positivo, pois se fosse negativo poderíamos trocar o sinal de u_1 ou de v_1 fazendo que ele virasse positivo, e isto iria contradizer o fato de ele ser máximo. Temos que $S^{n-1} \times S^{n-1}$ é uma variedade de dimensão $2(n-1)$, a função f é diferenciável e sua derivada num ponto (u, v) na direção do vetor tangente (X, Y) é dada por $df_{(u,v)}(X, Y) = MX \cdot v + Mu \cdot Y$. Como (u_1, v_1) é ponto de máximo de f, a derivada de f se anula em (u_1, v_1), portanto $MX \cdot v_1 + Mu_1 \cdot Y = 0$ para quaisquer $(X, Y) \in T_{(u_1,v_1)}S^{n-1} \times S^{n-1}$. Fazendo $Y = 0$

e depois $X = 0$, lembrando que o espaço tangente da esfera tem codimensão 1 em $\mathbb{R}^n$, existem $\alpha_1, \beta_1 \in \mathbb{R}$ tais que $Mu_1 = \alpha_1 v_1$ e $M^T v_1 = \beta_1 u_1$. Como o número $f(u_1, v_1) = Mu_1 \cdot v_1 \geq 0$, temos que $\alpha_1\beta_1 \geq 0$. É fácil ver que $MM^T u_1 = \alpha_1\beta_1 u_1$ e $M^T M v_1 = \alpha_1\beta_1 v_1$, portanto $(\alpha_1\beta_1)^{1/2}$ é valor singular de M. Denotemos $\sigma_1 = \alpha_1\beta_1$. Repetindo este método em $[u_1]^\perp \cap S^{n-1} \times [v_1]^\perp \cap S^{n-1}$, obtemos vetores u_2 e v_2, ortogonais a u_1 e v_1 respectivamente, e números reais $\alpha_2, \beta_2 > 0$ tais que $Mu_2 = \alpha_2 v_2$ e $M^T v_2 = \beta_2 u_2$, portanto $\sigma_2 = (\alpha_2\beta_2)^{1/2}$ é valor singular de M. Podemos repetir esse processo n vezes e obter uma lista $u_1, \ldots, u_n$ e $v_1, \ldots, v_n$ de vetores unitários em $\mathbb{R}^n$ com $u_i \cdot u_j = v_i \cdot v_j = 0$ para $i \neq j$, e uma lista de números reais $\alpha_3, \ldots, \alpha_n$ e $\beta_3, \ldots, \beta_n$ tais que $Mu_i = \alpha_i v_i$ e $M^T v_i = \beta_i u_i$, portanto $\sigma_i = (\alpha_i\beta_i)^{1/2}$ são valores singulares de M. Escrevendo cada um destes u e v em função da base canônica de $\mathbb{R}^n$ não é difícil ver que existem matrizes ortogonais U e V tais que $M = U\Sigma V^T$ onde Σ é uma matriz diagonal cujas diagonais são valores singulares de M. ■

Resultados gerais

Os resultados abaixo são bastante conhecidos e foram usados no texto. Incluímos as referências para estes fatos na bibliografia.

Lema 18 [4] *Dada uma função suave $f : X \to Y$ entre duas variedades X e Y de mesma dimensão e um ponto $x \in X$, se $df_x : TX_x \to TY_{f(x)}$ é isomorfismo, existem abertos $U \subset X$ e $V \subset Y$ tais que a restrição de f a U e V é um difeomorfismo.*

Lema 19 [4] *Dada uma função suave entre duas variedades X e Y de dimensão $m \geq n$, se $y \in Y$ é valor regular de f, temos que $Z = f^{-1}(y)$ é uma variedade de dimensão $m - n$ e $TZ_x = \ker(df_x : TX_x \to TY_{f(x)})$.*

Lema 20 [4] *Dadas uma função suave entre duas variedades X e Y de dimensão $m \geq n$, e um $x \in X$ tal que $df_x : TX_x \to TY_{f(x)}$ é sobrejetiva, temos que $f(x)$ está no interior de $f(X)$.*

Lema 21 [8] *Temos que* $\det(\exp(M)) = \exp(\mathrm{tr}(M))$. *Em particular, a exponencial de matrizes anti-simétricas $n \times n$ tem determinante 1.*

Lema 22 *O conjunto $\mathcal{S}'(n)$ das matrizes simétricas reais $n \times n$ com espectro simples é aberto no conjunto $\mathcal{S}(n)$ das matrizes simétricas reais $n \times n$. Além disso, a restrição da função espectro ordenado λ_o a $\mathcal{S}'(n)$ é suave.*

Prova: Primeiro vamos ver que $\mathcal{S}'(n)$ é um aberto em $\mathcal{S}(n)$ mostrando que seu complementar R, o conjunto das matrizes $n \times n$ reais simétricas com autovalores repetidos, é fechado em $\mathcal{S}(n)$. Usaremos a norma

$$||S|| = \max_{x \in S^{n-1}} Sx \cdot x = \max_{\lambda \text{ autovalor de } S} \lambda.$$

Seja $\{A_m\}_{m\in\mathbb{N}}$ uma sequência de matrizes em R convergindo para A_∞. Para cada m escolha λ_m um autovalor duplo de A_m. Pelo Teorema Espectral, existe um subespaço de $\mathbb{R}^n$ de dimensão 2 que é um auto-espaço para λ_m. Tome u_m e v_m dois vetores unitários e ortogonais em cada um destes subespaços.

Como $A_m \to A_\infty$ existe $M \in \mathbb{N}$ tal que se $m > M$ então

$$||A_m - A_\infty|| < 1 \ \ \Rightarrow ||A_m|| < ||A_\infty|| + 1.$$

Em outras palavras, se $m > M$ então os autovalores duplos λ_m estão contidos num compacto. Logo existe uma subsequência $\{\lambda_{m_\ell}\}$ convergente. Seja λ_∞ o limite desta subsequência. As sequências de vetores $\{u_m\}$ e $\{v_m\}$ estão contidas na esfera unitária de $\mathbb{R}^n$, logo também existem subsequências convergentes $\{u_{m_\ell}\}$ e $\{v_{m_\ell}\}$ convergindo para u_∞ e v_∞, respectivamente. Tomando índices de forma que as três subsequências convirjam, obtemos por continuidade:

$$\begin{aligned} &A_\infty u_\infty = \lambda_\infty u_\infty, \ \ A_\infty v_\infty = \lambda_\infty v_\infty, \\ &u_\infty \perp v_\infty, \ \ ||u_\infty|| = ||v_\infty|| = 1. \end{aligned}$$

O que mostra que λ_∞ é um autovalor duplo de A_∞ ($u_\infty \neq v_\infty$) e portanto R é fechado em $\mathcal{S}(n)$.

Considere um caminho suave $S : (-a, a) \to \mathcal{S}(n)$ com $S(0) \in \mathcal{S}'(n)$. Vamos mostrar que existe $b > 0$ tal que se $|t| < b$ então $S(t) \in \mathcal{S}'(n)$ e que $\lambda_o(S(t))$ varia suavemente com ϵ. Os autovalores (distintos) de $S(t)$ são as raízes de

$$p(\lambda) = p(t, \lambda) = \det(S(t) - \lambda I),$$

que é uma função suave (na verdade um polinômio) das entradas de $S(t)$, logo uma função suave em t. Como os autovalores de $S(0)$ são raízes simples de $p(0, \lambda)$, digamos $\lambda_1^0 < \lambda_2^0 < \ldots < \lambda_n^0$, temos que $p_\lambda(0, \lambda_i^0) \neq 0$, para $1 \leq i \leq n$. Pelo Teorema da Função Implícita existem $b_1, \ldots, b_n$ tais que $\lambda_i(t)$ dependem suavemente de t se $|t| < b_i$. Tome $b' = \min_{1\leq i\leq n} b_i$ e faça $\lambda_o(t) = (\lambda_1(t), \ldots, \lambda_n(t))$. Como o espectro de $S(0)$ é simples e $\lambda_i(t)$ são suaves se $|t| < b'$ existe $b < b'$ satisfazendo os pedidos. ∎

Referências Bibliográficas

[1] A. Horn, Doubly stochastic matrices and the diagonal of a rotation matrix, *Amer. J. Math.*, **76** (1954), 620–630.

[2] A. W. Marshall e I. Olkin, *Inequalities: theory of majorization and its applications*, Mathematics in Science and Engineering, **143**, Academic Press, New York, 1979.

[3] B. Kostant, Om convexity, the Weyl group and the Iwasawa decomposition, *Ann. Sci. École Norm. Sup.*, **6(4)**, (1973), 413–455.

[4] E. L. Lima, *Curso de Análise*, Vol. 2, IMPA, 2005.

[5] I. Schur, Über eine klase von mittelbindungen mit anwendungen auf der determinanten theorie, *S. B. Berlin Math. Ges.*, **22**, (1923), 9–20.

[6] J. Duistermaat, *The momentum maps*, Collection Topics in Differential Geometry Vol I e II, 347–392 (1988), Colloq. Math. Soc. Janos Bolyal, **46**.

[7] M. Atiyah, Convexity and commuting Hamiltonians, *Bull. London Math. Soc.*, **14**, (1982), 1–15.

[8] P. Lax, *Linear Algebra and its Applications*, John Wiley and Sons, Hoboken, NJ, 2007.

[9] R. S. Leite, *Uma demonstração geométrica do teorema de Schur-Horn*, Dissertação de Mestrado, Departamento de Matemática, PUC-Rio, 1996.

[10] R. S. Leite, T. R. W. Richa, C. Tomei, Geometric proofs of some theorems of Schur-Horn type, *Lin. Alg. App.*, **286**, (1999), 149–173.

[11] R. C. Thompson, Singular values, diagonal elements and convexity, *SIAM J. Appl. Math.*, **32(1)**, (1977), 39–63.

[12] T.-Y. Tam, A Lie theoretical approach of Thompson's theorems of singular values-diagonal elements and some related results, *J. of London Math. Soc.* **60**, (1999), 431–448.

[13] V. Guillemin e S. Sternberg, Convexity properties of the momentum mapping I, *Inv. Math.*, **67**, (1982), 491–513.