



Miguel Angel Orrillo Cumpa

**Análise em Grassmannianas e o Teorema de
Johnson-Lindenstrauss**

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-graduação em
Matemática da PUC-Rio como requisito parcial para obtenção
do grau de Mestre em Matemática.

Orientador: Prof. Carlos Tomei

Rio de Janeiro
Abril 2013



Miguel Angel Orrillo Cumpa

**Análise em Grassmannianas e o Teorema de
Johnson-Lindenstrauss**

Dissertação apresentada ao Programa de Pós-graduação em Matemática da PUC-Rio como requisito parcial para obtenção do grau de Mestre em Matemática. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Carlos Tomei

Orientador

Departamento de Matemática — PUC-Rio

Prof. Nicolau Corção Saldanha

Departamento de Matemática — PUC-Rio

Prof. Juliana Freire

Departamento de Matemática — PUC-Rio

Prof. Roberto Imbuzeiro M.F. de Oliveira

Instituto Nacional de Matemática Pura e Aplicada — IMPA

Prof. José Eugênio Leal

Coordenador Setorial do Centro Técnico Científico — PUC-Rio

Rio de Janeiro, 19 de abril de 2013

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Miguel Angel Orrillo Cumpa

Bacharel em Matemática Pura - Universidade Federal Fluminense.

Ficha Catalográfica

Orrillo, Miguel

Análise em Grassmannianas e o Teorema de Johnson-Lindenstrauss / Miguel Angel Orrillo Cumpa; orientador: Carlos Tomei. — Rio de Janeiro : PUC-Rio, Departamento de Matemática, 2013.

v., 63 f: il. ; 29,7 cm

1. Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Matemática.

Inclui referências bibliográficas.

1. Matemática – Dissertação. 2. Concentração de Medida. 3. Variedades de Grassmann. 4. Teorema de Johnson-Lindenstrauss. I. Tomei, Carlos. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Matemática. III. Título.

CDD: 510

A la memoria de Roxy Orrillo.

Agradecimentos

Benedictus Dominus Deus Israel. A Deus pela graça concedida desta dissertação. A Ele e todos os seus santos.

A Jaime Orrillo, meu professor da vida toda, por ter-me mostrado sempre a matemática de uma forma mais aprazível. A ele por sua infinita persistência. A mi madre Yolanda por su amor, dedicación, educación y por la tenacidad en darme siempre todo su apoyo. A mis abuelos José, Rosa y Laura, por sus oraciones y pedidos constantes de protección. A mis hermanas Leslie y Roxana (in memoriam) y a mi tía Juana por confiar en mi capacidad deseándome siempre lo mejor. A todos mis familiares en general, primos, primas, tíos y tías. A todos ellos por el cariño que siempre me dieron y que hicieron, aún estando lejos, sentir la confianza de pertenecer a una familia grande, unida y querida. Um agradecimento mais do que especial à minha companheira e esposa Amanda Orrillo pelo amor, pela compreensão e apoio que foram essências nesses momentos de dissertação. À minha querida sogra Serly Soares pelo papel de segunda mãe que ela desempenhou durante nossa convivência.

Ao meu orientador Carlos Tomei pela disponibilidade, entusiasmo e principalmente por me outorgar um pouco da sua erudição. Aos membros da banca: Nicolau Corção Saldanha, Roberto Imbuzeiro e Juliana Freire por aceitar e revisar a minha dissertação. Aos professores Jairo Bochi, Thomas Lewiner, Ricardo Sá Earp, Alex Castro, meus professores durante o mestrado na Puc-Rio. A eles muito obrigado pela formação.

À Pontifícia Universidade Católica do Rio de Janeiro pela bolsa de isenção e sua magnífica estrutura. Ao departamento de matemática da PUC-Rio e sua funcionária Creuza Nascimento pelo zelo e pelas constantes exortações.

Aos meus queridos amigos do peito, irmãos, Acir Carlos da Silva Junior, Johel Beltrán, João Marcos Breia Jucá, Juan Pablo Luna y Victor Goulart Nascimento pela leitura, correção ou sugestões outorgadas durante toda a realização desta dissertação. Aos meus queridos amigos Felipe Melo, Rodrigo Pacheco, Cong Zhou, Américo Cunha e André Zaccur pelas dicas de latex e gráficas. Um agradecimento especial ao meu querido amigo do peito Eric Biagioli pela organização "latexiana" da minha dissertação. Muito obrigado a todos os outros meus amigos que, felizmente sendo muitos, não acabaria de mencioná-los aqui. De uma maneira especial, ao meu querido padrinho espiritual Carlos Eduardo Guedes Belchior (Cadú). Um obrigado especialíssimo à minha querida Hazel Crato, ao meu amigo irmão de fé Juan Eduardo Casavilca, meu amigo Guillermo Gomez e à família Jucá (Débora, Celso, Pedro e Tche). A todos eles pela sincera torcida. Ao meu amigo Maycol junto com Carolina Parra (pelos '5' inegáveis) e seus familiares. Ao meu querido

”compadre” Daniel por sua amizade sincera e suas irreverências. A todos eles, um muito obrigado por fazer parte desta minha segunda família, meus amigos.

Finalmente, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa concedida.

Resumo

Orrillo, Miguel; Tomei, Carlos. **Análise em Grassmannianas e o Teorema de Johnson-Lindenstrauss**. Rio de Janeiro, 2013. 63p. Dissertação de Mestrado — Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

Seja V um conjunto de n pontos no espaço euclidiano X de dimensão d . Pelo teorema de Johnson-Lindenstrauss, existe uma projeção entre X e Y , outro espaço de dimensão k bastante menor, com a propriedade que as distâncias entre imagens de pontos de V sejam mantidas dentro de um fator c arbitrariamente próximo de 1. O teorema apresenta uma relação entre d , k e c , indicando a possibilidade de dramáticas reduções de dimensão para representações fidedignas de V . A demonstração emprega as Grassmannianas, as variedades de subespaços de dimensão k em X . São construídas cartas e uma medida homogênea em relação à ação natural do grupo ortogonal na Grassmanniana. O resultado segue estimando através de gaussianas certas integrais de caráter fortemente geométrico.

Palavras-chave

Concentração de Medida; Variedades de Grassmann; Teorema de Johnson-Lindenstrauss;

Abstract

Orrillo, Miguel; Tomei, Carlos (advisor) . **Grassmanian Analysis and the Johnson–Lindenstrauss Theorem.** Rio de Janeiro, 2013. 63p. M.Sc. Dissertation — Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

Let V be a set of n points in the Euclidean space X of dimension d . The Johnson-Lindenstrauss theorem states that there is a projection between X and Y , another Euclidean space of a smaller dimension k , with the property that images of points of X under projection do not differ by more than a multiplicative factor c arbitrarily close to 1. The theorem presents a relation among d , k and c , indicating the possibility of dramatic dimensional reduction of very faithful representations of V . The proof makes use of Grassmanians, the manifolds consisting of subspaces of dimension k in X . In the text, charts are presented, together with a measure which is homogeneous with respect to the natural action of the orthogonal group on the Grassmanian. The result follows by taking estimates using gaussians of certain integrals with a strong geometric flavor.

Keywords

Measure Concentration; Grassmann Manifolds; Johnson–Lindenstrauss Theorem;

Sumário

| | | |
|-----|--|----|
| 1 | Introdução | 12 |
| 2 | Preliminares | 15 |
| 3 | Concentração de Medida na Esfera | 18 |
| 3.1 | A função gama | 18 |
| 3.2 | Áreas de esferas e volumes de bolas | 19 |
| 3.3 | Coordenadas esféricas e o elemento de área de S^{d-1} | 21 |
| 3.4 | Concentração em faixas equatoriais | 24 |
| 4 | Análise em Grassmannianas | 28 |
| 4.1 | $G_k(\mathbb{R}^d)$ como subvariedade de $\mathcal{S}(\mathbb{R}^d)$ | 29 |
| 4.2 | A função posto | 30 |
| 4.3 | $G_k(\mathbb{R}^d)$ como variedade diferenciável | 37 |
| 5 | O Teorema de Johnson-Lindenstrauss | 43 |
| 5.1 | Medida de Haar em $\mathcal{O}(d)$ | 43 |
| 5.2 | Medidas invariantes em S^{d-1} e em $G_k(\mathbb{R}^d)$ | 46 |
| 5.3 | Gaussianas e o lema geométrico | 47 |
| 5.4 | Demonstração do teorema de Johnson-Lindenstrauss | 55 |
| A | Noções de Variedades | 58 |
| B | Noções de Probabilidade | 60 |
| | Referências Bibliográficas | 62 |

Lista de Figuras

| | | |
|-----|---|----|
| 1.1 | Faixa equatorial de espessura h subtendendo um ângulo α . | 12 |
| 1.2 | Área coberta por uma faixa de espessura h . | 12 |
| 1.3 | $C_1, C_2 \subset S^2$ representam os pontos da esfera com $ x_1 \geq t$. | 13 |
| 3.1 | Coordenadas esféricas em $d = 3$. | 22 |
| 3.2 | O gráfico representa a desigualdade $\cos(x) \leq \exp(-x^2/2)$. | 26 |

Mesmo que eu tivesse o dom da profecia, e conhecesse todos os mistérios e toda a ciência; mesmo que tivesse toda a fé, a ponto de transportar montanhas, se não tiver caridade, não sou nada.

1 Coríntios, 13:2.

1

Introdução

Suponha que você queira pintar a superfície da Terra de raio ρ . Para isso você conta com a ajuda de um rolo capaz de pintar faixas gigantes ao redor da Terra. Que fração x da área da esfera você terá coberto depois de passar o rolo ao redor do equador pintando assim uma faixa de espessura h subtendendo um ângulo α ?

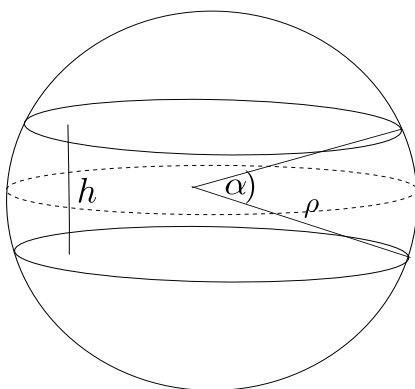


Figura 1.1: Faixa equatorial de espessura h subtendendo um ângulo α .

O problema é resolvido usando o seguinte fato geral:

Na esfera de raio ρ , a área de uma faixa qualquer de espessura h é $2\pi\rho h$.

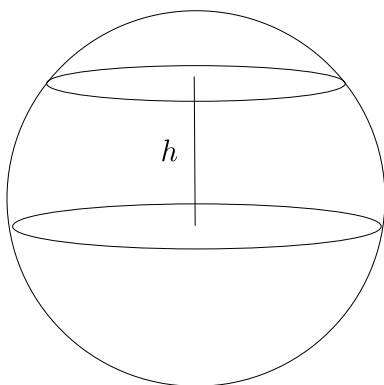


Figura 1.2: Área coberta por uma faixa de espessura h .

Dado que faixa tem espessura h e subtende um ângulo α , então $\text{sen}(\alpha/2) = h/2\rho$. Logo, uma simples regra de três permite calcular a porcentagem

$$x = \frac{2\pi\rho h}{4\pi\rho^2} 100 = \frac{h}{2\rho} 100 = \text{sen}\left(\frac{\alpha}{2}\right) 100.$$

Por exemplo, se $\alpha = 10^\circ$ (cuja faixa tem mais de 1 000 Km de largura), então a fração pintada seria aproximadamente 8,7%.

O fenômeno de concentração de medida garante que sua tarefa seria muito mais simples se a Terra fosse de dimensão alta. Mais especificamente, uma só volta à Terra com o rolo ao redor do equador seria suficiente para cobrir praticamente toda sua superfície. Isto significa que, para d muito grande, a área da superfície da esfera $S^{d-1} = \{x \in \mathbb{R}^d : \|x\| = 1\}$ é concentrada em torno do equador. Ou seja, a área do complemento da faixa equatorial (duas calotas esféricas) tende a 0 quando d tende ao infinito.

O teorema de Johnson-Lindenstrauss é o resultado principal desta dissertação. Ele mostra que a geometria de um conjunto V com n pontos não é muito perturbada por certas projeções ortogonais sobre subespaços de dimensão logarítmica de n . Em outras palavras, é possível projetar V em subespaços de dimensão baixa preservando bem a distância entre eles.

A prova usa um lema geométrico relacionado com a concentração da área em faixas equatoriais. Tome $s = (x_1, \dots, x_d)$ um ponto de S^{d-1} escolhido uniformemente e seja $\pi_{\mathbb{R}^k}(s) = (x_1, \dots, x_k)$ a projeção de s em suas k primeiras coordenadas –projeção ortogonal em \mathbb{R}^k –. Considere β um número próximo de 1 de modo que $t = \sqrt{\beta \frac{k}{d}} < 1$. Qual seria a probabilidade de $\|\pi_{\mathbb{R}^k}(s)\|$ ser maior ou igual que t ? Quando $k = 1$ e $d = 3$, a geometria fica clara na seguinte figura.

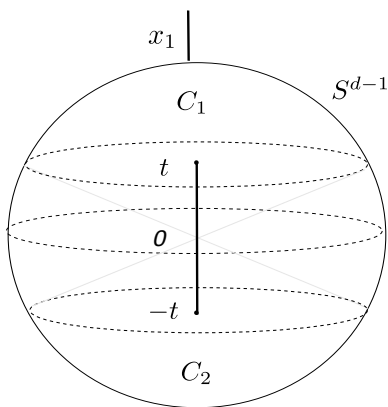


Figura 1.3: $C_1, C_2 \subset S^2$ representam os pontos da esfera com $|x_1| \geq t$.

Observamos que $\Pr \left[|x_1| > \sqrt{\beta \frac{k}{d}} \right] = 1 - x$, onde x é a fração de área da

esfera na faixa de espessura t . Quando d tende ao infinito, pela concentração de medida em faixas equatoriais, concluiríamos que $\Pr [|x_1| > t]$ tende a 0.

Um resultado mais geral vale para $k < d$ qualquer – para uma melhor descrição, o leitor pode consultar o Lema 5.6 do Capítulo 5.– Essencialmente, o lema afirma que o quadrado da norma da projeção em \mathbb{R}^k de pontos da esfera S^{d-1} , escolhidos uniformemente, se encontra fortemente concentrado em torno de sua média k/d .

A afirmação acima tem outra interpretação: em vez de variar o ponto e fixar uma projeção ortogonal de imagem com dimensão k , podemos fixar o ponto e fazer a média sobre todas as projeções ortogonais possíveis com imagem de dimensão k . A prova do teorema emprega o fato de que a norma de ambas as projeções possuem a mesma distribuição.

2 Preliminares

Neste capítulo coletaremos alguns fatos básicos. Todas as definições e resultados aqui apresentados são bem conhecidos e podem ser encontrados na literatura.

Supomos que o leitor tenha familiaridade com os conceitos de espaços topológicos, métricos e de medida, grupos e subgrupos e alguns conceitos e resultados básicos de álgebra linear, tais como projeções ortogonais, o teorema espectral, processo de ortogonalização de Gram-Schmidt, etc. Para o vocabulário básico sobre variedades diferenciáveis e de probabilidade, o leitor pode considerar os apêndices.

Alguns conjuntos de matrizes serão usados extensivamente. Seja $\mathcal{M}(d, \mathbb{R})$ o conjunto de todas as matrizes reais $d \times d$ com a norma de Frobenius

$$\|A\|_2 = \left(\sum_{i=1}^d \sum_{j=1}^d a_{ij}^2 \right)^{1/2}, \quad A = (a_{ij}).$$

O *grupo linear* formado por todas as matrizes inversíveis $d \times d$

$$GL(\mathbb{R}^d) = \{A \in \mathcal{M}(d, \mathbb{R}) : \det(A) \neq 0\}.$$

é um subconjunto aberto de $\mathcal{M}(d, \mathbb{R})$. O *grupo ortogonal* $\mathcal{O}(d)$ é um subgrupo de $GL(\mathbb{R}^d)$, definido como

$$\mathcal{O}(d) = \{Q \in GL(\mathbb{R}^d) : Q^*Q = QQ^* = I\}.$$

Além disso, $\mathcal{O}(d)$ é uma subvariedade compacta suave (C^∞) de dimensão $d(d-1)/2$.

A esfera $S^{d-1} = \{x \in \mathbb{R}^d : \|x\|^2 = \sum_{i=1}^d x_i^2 = 1\}$ é uma subvariedade compacta suave de codimensão 1 e classe C^∞ .

Um grupo G é um *grupo topológico* quando G é um espaço topológico e as operações de multiplicação e de tomada de inverso, $(a, b) \mapsto ab$ e $a \mapsto a^{-1}$, são contínuas.

Seja M uma variedade diferenciável e G um grupo topológico. Uma *ação*

de G sobre M é uma aplicação contínua $\phi : G \times M \rightarrow M$ tal que

- (i) Se e é o elemento neutro de G então $\phi(e, x) = x, \forall x \in M$
- (ii) $\phi(g, \phi(h, x)) = \phi(gh, x)$ para todo $g, h \in G$ e para todo $x \in M$.

A ação $\phi : G \times M \rightarrow M$ é *transitiva*, ou G age *transitivamente* em M através de ϕ , se para todo $x, y \in M$, existe $g \in G$ tal que $\phi(g, x) = y$. Nesse texto, vamos usar a ação do grupo ortogonal $G = \mathcal{O}(d)$ sobre a esfera $M = S^{d-1}$, dada por

$$\begin{aligned} \phi : \mathcal{O}(d) \times S^{d-1} &\rightarrow S^{d-1} \\ (g, x) &\mapsto gx. \end{aligned}$$

Seja ϕ uma ação de um grupo G sobre uma variedade M e $x \in M$. A ϕ -órbita de x é por definição o conjunto $Gx = \{\phi(g, x) : g \in G\}$.

A ação $\phi : G \times M \rightarrow M$ é *própria* se e somente se $\tilde{\phi} : G \times M \rightarrow M \times M$ definida por $\tilde{\phi}(g, x) = (x, \phi(g, x))$ é uma aplicação própria, isto é, inversa de compacto é compacto. A demonstração da seguinte proposição pode ser encontrada em (1, p. 265).

Proposição 2.0.1 *Seja $\phi : G \times M \rightarrow M$ uma ação própria. Então toda órbita Gx é uma subvariedade fechada de M .*

Lembramos ao leitor que em um espaço topológico X a σ -álgebra de Borel é a menor σ -álgebra que contém todos os abertos de X . Seus elementos são os *borelianos* de X .

Funções transferem medidas de um espaço para outro através do *push-forward*. Mais especificamente, sejam (X, \mathfrak{X}, μ) um espaço de medida, Y um espaço topológico qualquer e $\phi : X \rightarrow Y$ uma aplicação. A *medida induzida* $\phi_*\mu$ é definida para todo boreliano B de Y como

$$\phi_*\mu(B) := \mu(\phi^{-1}(B)) = \mu(\{g \in G : \phi(g) \in B\}). \quad (2-1)$$

Uma medida ν de M é *invariante* pela ação $\phi : G \times M \rightarrow M$ se para todo $g \in G$ e para todo conjunto boreliano $A \subset M$, vale

$$\nu(A) = \nu(gA) = \nu(\{\phi(g, a) : a \in A\}).$$

Sejam $\mathcal{B}(\mathbb{R}^d)$ a σ -álgebra de Borel em \mathbb{R}^d e λ a medida de Lebesgue. Uma ação do grupo ortogonal $\mathcal{O}(d)$ sobre \mathbb{R}^d é naturalmente dada por

$$\phi : \mathcal{O}(d) \times \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad \phi(g, x) = gx. \quad (2-2)$$

Terminamos o capítulo mostrando que a medida de Lebesgue é invariante por essa ação.

Lembramos ao leitor que toda transformação linear $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ é o produto de *transformações elementares*, dadas por

$$T_1(x_1, \dots, x_i, \dots, x_j, \dots, x_d) = (x_1, \dots, x_j, \dots, x_i, \dots, x_d)$$

$$T_2(x_1, \dots, x_i, \dots, x_d) = (x_1, \dots, cx_i, \dots, x_d) \quad (c \neq 0)$$

$$T_3(x_1, \dots, x_i, \dots, x_k, \dots, x_d) = (x_1, \dots, x_i + cx_k, \dots, x_k, \dots, x_d) \quad (k \neq i).$$

Proposição 2.0.2 *Seja $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ uma transformação linear inversível e $A \subset \mathbb{R}^d$ um boreliano. Então*

$$\lambda(T(A)) = |\det(T)| \lambda(A). \quad (2-3)$$

Demonstração: Suponha que a proposição seja verdadeira para transformações T e S . Então ela vale para $T \circ S$

$$\lambda((T \circ S)(A)) = |\det(T)| \lambda(S(A)) = |\det(T)| |\det(S)| \lambda(A) = |\det(T \circ S)| \lambda(A).$$

Logo, é suficiente mostrar que a proposição é válida para as transformações elementares T_1 , T_2 e T_3 . Explicitamos o argumento para T_3 onde $\det(T_3) = 1$. Usamos Fubini afim de trocar a ordem de integração entre x_i e x_j .

$$\begin{aligned} |\det(T_3)| \lambda(A) &= \lambda(A) = \int_A 1 \, dx_1 \cdots dx_i \cdots dx_j \cdots dx_d \\ &= \int_A 1 \, dx_1 \cdots dx_j \cdots dx_i \cdots dx_d \\ &= \int_{T_3(A)} 1 \, dx_1 \cdots dx_i \cdots dx_j \cdots dx_d = \lambda(T_3(A)). \end{aligned}$$

□

Corolário 2.1 *A medida de Lebesgue é invariante pela ação do grupo ortogonal.*

Demonstração: Transformações lineares $g \in \mathcal{O}(d)$ satisfazem $gg^* = I$, o que implica que $|\det(g)| = 1$. Logo, $\lambda(g(A)) = \lambda(A)$. □

3

Concentração de Medida na Esfera

Um primeiro exemplo do fenômeno de concentração de medida ocorre na esfera $S^{d-1} \subset \mathbb{R}^d$. Este capítulo tem o propósito de apresentá-lo detalhadamente: a medida uniforme da esfera se concentra em faixas equatoriais quando a dimensão tende ao infinito — para uma descrição precisa, o leitor pode consultar a equação (3-5).

Começamos introduzindo a função gama e algumas contas básicas: o cálculo das áreas das esferas e do volume das bolas na métrica euclidiana em \mathbb{R}^d , cujas fórmulas envolvem a função gama. Finalmente, mostraremos a concentração da esfera em faixas equatoriais. Aqui, usaremos coordenadas esféricas para calcular a área de calotas esféricas.

3.1

A função gama

A *função gama* é

$$\Gamma(z) = \int_0^{\infty} t^{z-1} \exp(-t) dt, \quad \text{para } z \in \mathbb{C}, \operatorname{Re}(z) > 0.$$

Para conveniência do leitor, mostramos sua integrabilidade.

Proposição 3.1.1 *A função gama é bem definida.*

Demonstração: O argumento é diferente para t perto e longe de 0

$$\int_0^{\infty} |t^{z-1} \exp(-t)| dt = \int_0^1 |t^{z-1} \exp(-t)| dt + \int_1^{\infty} |t^{z-1} \exp(-t)| dt.$$

Para a primeira parcela, como $t \geq 0$, temos $\exp(-t) \leq 1$ e assim

$$\begin{aligned} |t^{z-1} \exp(-t)| &\leq |t^{z-1}| = |\exp((z-1) \log t)| \\ &= \exp((\operatorname{Re}(z) - 1) \log t) = \exp(\log t^{\operatorname{Re}(z)-1}) = t^{(\operatorname{Re}(z)-1)}. \end{aligned}$$

Como $\int_0^1 t^a < \infty$ para todo $a > -1$,

$$\int_0^1 |t^{z-1} \exp(-t)| dt \leq \int_0^1 t^{\operatorname{Re}(z)-1} < \infty.$$

Para a segunda parcela, $t \geq 1$ e temos

$$|t^{z-1} \exp(-t)| = t^{\operatorname{Re}(z)-1} \exp(-t).$$

A função $h(t) = t^{\operatorname{Re}(z)-1} \exp(-t/2)$ é contínua em $[1, \infty)$ e $\lim_{t \rightarrow \infty} h(t) = 0$, logo h é limitada, $|h| \leq C_z$. Assim,

$$|t^{z-1} \exp(-t)| = t^{\operatorname{Re}(z)-1} \exp(-t) \leq C_z \exp(-t/2).$$

Como $\int_1^\infty C_z \exp(-t/2) < \infty$, a segunda parcela também é integrável. \square

Integrando por partes,

$$\int_0^\infty t^z \exp(-t) dt = -t^z \exp(-t) \Big|_0^\infty + z \int_0^\infty t^{z-1} \exp(-t) dt = z \int_0^\infty t^{z-1} \exp(-t) dt,$$

concluimos que para $\operatorname{Re}(z) > 0$ a função gama satisfaz a conhecida equação funcional $\Gamma(z+1) = z\Gamma(z)$. Isto permite estendê-la para quase todos os números complexos. Além disso, por indução,

$$\Gamma(n+1) = n! \quad \text{e} \quad \Gamma\left(k + \frac{1}{2}\right) = \frac{\sqrt{\pi} (2k-1)!!}{2^k}. \quad (3-1)$$

O fatorial duplo $(2k-1)!!$ é dado pelo produto de todos os naturais ímpares até $(2k-1)$.

3.2

Áreas de esferas e volumes de bolas

Indicamos por $S_a^{d-1}(R)$ e $B_a^d(R)$ respectivamente a esfera e a bola em \mathbb{R}^d de raio R centradas em $a \in \mathbb{R}^d$,

$$S_a^{d-1}(R) = \{x \in \mathbb{R}^d : \|x - a\| = R\}, \quad B_a^d(R) = \{x \in \mathbb{R}^d : \|x - a\| \leq R\}.$$

Quando $R = 1$ e $a = 0$, escreveremos simplesmente S^{d-1} e B^d .

Agora definiremos a medida de Lebesgue na esfera S^{d-1} . Lembramos ao leitor que os borelianos da esfera $S_0^{d-1}(R)$ são todas as interseções da própria esfera com os borelianos de \mathbb{R}^d . Se $A \subset S_0^{d-1}(R)$ é um boreliano da esfera, denotemos por $C(A)$ o conjunto mensurável de \mathbb{R}^d dado por $C(A) = \{rx : r \in [0, 1], x \in A\}$. Definimos, então, a *medida de Lebesgue* $\eta(A)$ de $A \in \mathcal{B}(S_R^{d-1})$ como sendo

$$\eta(A) = \frac{d}{R} \lambda(C(A)). \quad (3-2)$$

Fazendo $R = 1$, $(S^{d-1}, \mathcal{B}(S^{d-1}), \eta)$ é um espaço de medida.

Proposição 3.2.1 *A medida de Lebesgue η em S^{d-1} é invariante pela ação do grupo ortogonal $\mathcal{O}(d)$.*

Demonstração: Dado que a medida de Lebesgue em \mathbb{R}^d é invariante pela ação do grupo $\mathcal{O}(d)$,¹ a demonstração é direta. Observamos primeiro que se $g \in \mathcal{O}(d)$ e $A \in \mathcal{B}(S^{d-1})$ então certamente $gC(A) = C(gA)$. Assim,

$$\eta(A) = d \lambda(C(A)) = d \lambda(gC(A)) = d \lambda(C(gA)) = \eta(gA). \quad \square$$

A medida η definida em $\mathcal{B}(S^{d-1})$ é o análogo à medida que generaliza o comprimento em S^1 assim como λ é o análogo a área em \mathbb{R}^2 . Em \mathbb{R}^3 , temos $S^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$ como fronteira de $B^3 = \{x \in \mathbb{R}^3 : \|x\| \leq 1\}$. Neste caso, η é o que conhecemos por área (na esfera) e λ como volume.

Proposição 3.2.2

$$\lambda(B_0^d(R)) = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2} + 1)} R^d, \quad \eta(S_0^{d-1}(R)) = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} R^{d-1}.$$

Demonstração: Primeiro mostraremos que

$$\lambda(B_0^d(R)) = \begin{cases} \frac{\pi^{d/2}}{(\frac{d}{2})!} R^d & , \text{ se } d \text{ é par;} \\ \frac{2^{(d+1)/2} \pi^{(d-1)/2}}{d!!} R^d & , \text{ se } d \text{ é ímpar.} \end{cases} \quad (3-3)$$

Denotemos como $V_d(R) = \lambda(B_0^d(R))$. A geometria fica mais clara para $d = 3$. Usando coordenadas polares, integramos no disco de raio R , onde para cada ponto p temos uma bola de dimensão 1 (intervalo) de centro p . Assim,

$$V_3(R) = \int_0^R \int_0^{2\pi} V_1(\sqrt{R^2 - r^2}) r d\theta dr = \int_0^{2\pi} \int_0^R (2\sqrt{R^2 - r^2}) r dr d\theta$$

e, fazendo a mudança de variável $u = R^2 - r^2$,

$$V_3(R) = \int_0^{2\pi} 1 d\theta \int_{R^2}^0 -u^{1/2} du = 2\pi \int_0^{R^2} u^{1/2} du = 2\pi \left(\frac{2u^{3/2}}{3} \Big|_0^{R^2} \right) = \frac{4\pi}{3} R^3.$$

Seja $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $T(x) = Rx$. Então aplicando a Proposição 2.0.2, é fácil calcular o volume de uma bola de raio R a partir da bola unitária:

¹Veja o Corolário 2.1 no capítulo 2.

$V_d(R) = R^d V_d(1)$. Assim, obtemos a seguinte relação de recorrência

$$\begin{aligned} V_d(R) &= \int_0^R \int_0^{2\pi} V_{d-2}(\sqrt{R^2 - r^2}) r d\theta dr \\ &= \int_0^{2\pi} \int_0^R (R^2 - r^2)^{(d-2)/2} V_{d-2}(1) r dr d\theta = \pi V_{d-2}(1) \int_0^{R^2} u^{(d-2)/2} du \\ &= 2\pi V_{d-2}(1) \frac{u^{d/2}}{d} \Big|_0^{R^2} = 2\pi V_{d-2}(1) \frac{R^d}{d} = \frac{2\pi R^2}{d} V_{d-2}(R). \end{aligned}$$

Agora usamos indução em d para verificar o caso geral. Para $d = 1, 2$ temos $V_1(R) = 2R$ e $V_2(R) = \pi R^2$, de acordo com a fórmula (3-3).

Supondo a equação (3-3) válida para d par, vamos verificar a sua validade para $d + 2$ usando a fórmula $V_d(R) = \frac{2\pi R^2}{d} V_{d-2}(R)$

$$V_{d+2}(R) = \frac{2\pi R^2}{d+2} V_d(R) = \frac{2\pi R^2}{d+2} \frac{\pi^{d/2}}{(\frac{d}{2})!} R^d = \frac{2\pi^{(d+2)/2}}{2 \frac{(d+2)}{2} (\frac{d}{2})!} = \frac{\pi^{(d+2)/2}}{(\frac{d+2}{2})!}.$$

Podemos fazer o mesmo para d ímpar

$$V_{d+2}(R) = \frac{2\pi R^2}{d+2} V_d(R) = \frac{2\pi R^2}{d+2} \frac{2^{(d+1)/2} \pi^{(d-1)/2}}{(d)!!} R^d = \frac{2^{((d+2)+1)/2} \pi^{((d+2)-1)/2}}{(d+2)!!} R^{d+2}.$$

Agora verificaremos a fórmula envolvendo a função gama. Com efeito, para $d = 2k$,

$$\frac{\pi^{2k/2}}{\Gamma(\frac{2k}{2} + 1)} R^{2k} = \frac{\pi^k}{\Gamma(k + 1)} R^{2k} = \frac{\pi^k}{k!} R^{2k} = V_{2k}(R)$$

e para $d = 2k + 1$,

$$\frac{\pi^{(2k+1)/2}}{\Gamma(\frac{2k+1}{2} + 1)} R^{2k+1} = \frac{\pi^{(2k+1)/2}}{2 \frac{2k+1}{2} \Gamma(k + \frac{1}{2})} R^{2k+1} = \frac{\pi^k 2^{k+1}}{(2k + 1)!!} R^{2k+1} = V_{2k+1}(R).$$

Assim, concluímos que $\lambda(B_0^d(R)) = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2} + 1)} R^d$, para todo d . Da definição, $\eta(A) = \frac{d}{R} \lambda(C(A))$, concluímos também

$$\eta(S_0^{d-1}(R)) = \frac{d}{R} V_d(R) = \frac{d}{R} \frac{\pi^{d/2}}{\Gamma(\frac{d}{2} + 1)} R^d = d \frac{\pi^{d/2}}{\frac{d}{2} \Gamma(\frac{d}{2})} R^{d-1} = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})} R^{d-1}. \quad \square$$

3.3

Coordenadas esféricas e o elemento de área de S^{d-1}

Sejam $U, V \subset \mathbb{R}^d$ abertos, $\psi : U \rightarrow V$ um difeomorfismo C^1 e $X \subset U$ mensurável com $\psi(X) \subset V$ também mensurável. Considere $f : \psi(X) \rightarrow \mathbb{R}$

uma função integrável. Pelo teorema da mudança de variáveis,

$$\int_{\psi(X)} f(y) dy = \int_X f(\psi(x)) \cdot |\det D\psi(x)| dx.$$

A função ψ é o que chamamos de mudança de variável. Veja (10). Entre os exemplos estão as mudanças de coordenadas tradicionais. As *coordenadas esféricas* consistem de uma coordenada radial $\rho \geq 0$ e ângulos $\phi_1, \phi_2, \dots, \phi_{d-2} \in [0, \pi]$ e $\phi_{d-1} \in [0, 2\pi)$, definidos pelas fórmulas

$$\begin{aligned} x_1 &= \rho \cos(\phi_1) \\ x_2 &= \rho \operatorname{sen}(\phi_1) \cos(\phi_2) \\ x_3 &= \rho \operatorname{sen}(\phi_1) \operatorname{sen}(\phi_2) \cos(\phi_3) \\ &\vdots \\ x_{d-1} &= \rho \operatorname{sen}(\phi_1) \cdots \operatorname{sen}(\phi_{d-2}) \cos(\phi_{d-1}) \\ x_d &= \rho \operatorname{sen}(\phi_1) \cdots \operatorname{sen}(\phi_{d-2}) \operatorname{sen}(\phi_{d-1}). \end{aligned}$$

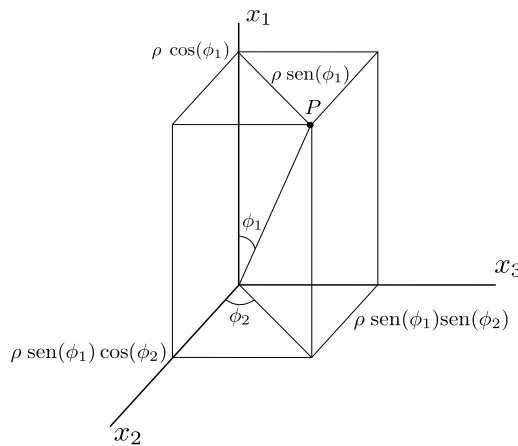


Figura 3.1: Coordenadas esféricas em $d = 3$.

No aberto $U = (0, \infty) \times (0, \pi) \times (0, \pi), \times \dots, \times (0, 2\pi)$, a função

$$\begin{aligned} \psi_d : U &\rightarrow \psi_d(U) \\ (\rho, \phi_1, \dots, \phi_{d-1}) &\mapsto (x_1, x_2, \dots, x_d) \end{aligned}$$

é um difeomorfismo (de fato, $\rho = \|x\|$ e os ângulos são obtidos sequencialmente). Os pontos de \mathbb{R}^d fora de $\psi(U)$ têm medida zero, e não são relevantes para o cálculo das integrais.

Sejam $X \subset U$ e $\psi_d(X)$ mensuráveis. Mudando de variáveis, temos

$$\lambda(\psi_d(X)) = \int_{\psi_d(X)} dx = \int_X |\det D\psi_d((\rho, \phi_1, \dots, \phi_{d-1}))| d\rho d\phi_1 \cdots d\phi_{d-1}.$$

Proposição 3.3.1 Para $d \geq 3$ e para $(\rho, \phi_1, \dots, \phi_{d-1}) \in U$

$$\det D\psi_d((\rho, \phi_1, \dots, \phi_{d-1})) = \rho^{d-1} \text{sen}^{d-2}(\phi_1) \text{sen}^{d-3}(\phi_2) \cdots \text{sen}(\phi_{d-2}) > 0.$$

O caso $d = 2$ é a troca de variáveis habitual para coordenadas polares: a fórmula vale com uma interpretação adequada.

Demonstração: A prova é uma indução em d . Escrevemos $c_i = \cos(\phi_i)$, $s_i = \text{sen}(\phi_i)$ e $\det D\psi_d = \det D\psi_d(\rho, \phi_1, \dots, \phi_{d-1})$. Começamos mostrando o caso $d = 3$:

$$\begin{aligned} \det D\psi_3 &= \det \begin{vmatrix} c_1 & -\rho s_1 & 0 \\ s_1 c_2 & \rho c_1 c_2 & -\rho s_1 s_2 \\ s_1 s_2 & \rho c_1 s_2 & \rho s_1 c_2 \end{vmatrix} \\ &= \rho s_1 c_2 \det \begin{vmatrix} c_1 & -\rho s_1 \\ s_1 c_2 & \rho c_1 c_2 \end{vmatrix} + \rho s_1 s_2 \det \begin{vmatrix} c_1 & -\rho s_1 \\ s_1 s_2 & \rho c_1 s_2 \end{vmatrix} \\ &= \rho c_1 c_2 [\rho c_1^2 c_2 + \rho s_1^2 c_2] + \rho s_1 s_2 [c_1^2 \rho s_2 + \rho s_1^2 s_2] \\ &= \rho^2 s_1 c^2 + \rho^2 s_1 s_2^2 = \rho^2 \text{sen}(\phi_1). \end{aligned}$$

Note que as duas matrizes 2×2 acima são obtidas multiplicando a segunda linha de $D\psi_2$ por c_2 e por s_2 respectivamente: a indução explora esse padrão.

A hipótese de indução para $d = k$ afirma que

$$\begin{aligned} \det D\psi_k &= \det \begin{vmatrix} c_1 & -\rho s_1 & \cdots & 0 \\ s_1 c_2 & \rho c_1 c_2 & \cdots & 0 \\ s_1 s_2 c_3 & \rho c_1 s_2 c_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ s_1 \cdots c_{k-1} & \rho c_1 \cdots c_{k-1} & \cdots & -\rho s_1 \cdots s_{k-1} \\ s_1 \cdots s_{k-1} & \rho c_1 \cdots s_{k-1} & \cdots & \rho s_1 \cdots c_{k-1} \end{vmatrix} \\ &= \rho^{k-1} \text{sen}^{k-2}(\phi_1) \text{sen}^{k-3}(\phi_2) \cdots \text{sen}(\phi_{k-2}). \end{aligned}$$

Para $d = k + 1$, temos

$$\det D\psi_{k+1} = \det \begin{vmatrix} & E_1 D\psi_k & \alpha e_k \\ \hline s_k e_k^\top D\psi_k & & \beta \end{vmatrix}$$

onde e_k é o k -ésimo vetor canônico de \mathbb{R}^k , $\alpha = -\rho s_1 \cdots s_{k-1} s_k$, $E_1 = \text{diag}(1, \dots, 1, c_k)$ e $\beta = \rho s_1 \cdots s_{k-1} c_k$. Assim, concluímos a indução

$$\begin{aligned} \det D\psi_{k+1} &= \beta \det(E_1 D\psi_k) - \alpha s_k \det(D\psi_k) \\ &= \rho s_1 \cdots s_{k-1} c_k c_k \det D\psi_k + \rho s_1 \cdots s_{k-1} s_k s_k \det D\psi_k \\ &= \rho s_1 \cdots s_{k-1} \det D\psi_k \\ &= \rho^k \text{sen}^{k-1}(\phi_1) \text{sen}^{k-2}(\phi_2) \cdots \text{sen}^2(\phi_{k-2}) \text{sen}(\phi_{k-1}). \end{aligned}$$

Como $\rho \in (0, \infty)$, $\phi_i \in (0, \pi)$ para $i = 1, \dots, d-2$, então

$$\det D\psi_d((\rho, \phi_1, \dots, \phi_{d-1})) = \rho^{d-1} \text{sen}^{d-2}(\phi_1) \text{sen}^{d-3}(\phi_2) \cdots \text{sen}(\phi_{d-2}) > 0. \quad \square$$

Assim,

$$\lambda(\psi_d(X)) = \int_X \rho^{d-1} \text{sen}^{d-2}(\phi_1) \text{sen}^{d-3}(\phi_2) \cdots \text{sen}(\phi_{d-2}) d\phi_1 \cdots d\phi_{d-1}.$$

Lembramos ao leitor que $\lambda = m \times \eta$, onde m é uma medida em $(0, \infty)$ definida por $m(C) = \int_C \rho^{d-1} d\rho$, para todo boreliano $C \subset (0, \infty)$. Logo, para um boreliano $A \in S^{d-1}$,

$$\eta(A) = \int_{\psi_d^{-1}(A)} \text{sen}^{d-2}(\phi_1) \text{sen}^{d-3}(\phi_2) \cdots \text{sen}(\phi_{d-2}) d\phi_1 \cdots d\phi_{d-1}.$$

3.4 Concentração em faixas equatoriais

Vamos provar a concentração da esfera em faixas equatoriais. Para isso escolhemos um equador padrão, faixas equatoriais de espessura fixa e calotas. Nosso trabalho se resumirá a estimar áreas de calotas esféricas.

O *equador padrão* (ou simplesmente equador) em S^{d-1} é o conjunto

$$\begin{aligned} E^{d-1} &= \{(x_1, \dots, x_d) \in \mathbb{R}^d : x_1 = 0\} \\ &= \psi_d(\{(\rho, \phi_1, \dots, \phi_{d-1}) : \rho = 1, \phi_1 = \pi/2\}). \end{aligned}$$

Uma *faixa equatorial* $E_\epsilon^{d-1} \subset S^{d-1}$ de espessura 2ϵ com $\epsilon \in (0, 1)$ é

$$\begin{aligned} E_\epsilon^{d-1} &= \{(x_1, \dots, x_d) \in \mathbb{R}^d : -\epsilon < x_1 < \epsilon\} \\ &= \psi_d(\{(\rho, \phi_1, \dots, \phi_{d-1}) : \rho = 1, \pi/2 - \theta < \phi_1 < \pi/2 + \theta\}) \end{aligned}$$

onde $\theta = \text{arc sen}(\epsilon)$.

A medida de probabilidade uniforme para os conjuntos borelianos da esfera é definida por normalização

$$\sigma(A) = \frac{\eta(A)}{\eta(S^{d-1})}. \quad (3-4)$$

Novamente, a invariância da medida de Lebesgue implica que σ é invariante pela ação do grupo ortogonal $\mathcal{O}(d)$ ². A concentração da medida uniforme da esfera em faixas equatoriais de espessura 2ϵ se resume ao limite

$$\lim_{d \rightarrow \infty} \sigma(S^{d-1} \setminus E_\epsilon^{d-1}) = 0. \quad (3-5)$$

Claramente, $S^{d-1} \setminus E_\epsilon^{d-1} = \{(x_1, \dots, x_d) \in \mathbb{R}^d : \epsilon < |x_1| < 1\}$ é união de duas *calotas esférica*: uma calota é o subconjunto conexo da esfera acima ou abaixo de um plano $x_1 = c$ com $-1 < c < 1$. Mais especificamente

$$C_N^{d-1}(c) = \{x \in S^{d-1} : c < x_1 < 1\}; \quad C_S^{d-1}(c) = \{x \in S^{d-1} : -1 < x_1 < c\}.$$

Logo

$$\sigma(S^{d-1} \setminus E_\epsilon^{d-1}) = \sigma(C_N^{d-1}(\epsilon)) + \sigma(C_S^{d-1}(-\epsilon)) = 2 \sigma(C_N^{d-1}(\epsilon)).$$

Vamos ver que a medida de calotas esféricas $C_N^d(\epsilon)$ de dimensão d vai a zero quando $d \rightarrow \infty$. Em coordenadas esféricas,

$$C_N^{d-1}(\epsilon) = \psi_d\{(\rho, \phi_1, \dots, \phi_{d-1}) : \rho = 1, 0 < \phi < \pi/2 - \theta\}.$$

Precisamos de uma estimativa.

Proposição 3.4.1 Para $0 \leq \alpha \leq \frac{\pi}{2}$, vale $\cos^d(x) \leq \cos(x) \leq \exp(-x^2/2)$.

Demonstração: A cota $\cos^d(x) \leq \cos(x)$ é óbvia. A fim de mostrar a outra, definimos

$$f(x) = \ln(\cos(x)) + \frac{x^2}{2}, \quad x \in [0, \pi/2),$$

cuja derivada é $f'(x) = -\tan(x) + x$. Como $x < \tan(x)$ para $x \in [0, \pi/2)$, concluímos que $f'(x) < 0$ no intervalo $[0, \pi/2)$. Logo, $f(x) < f(0) \forall x \in [0, \pi/2)$, ou seja, $\ln(\cos(x)) < \frac{-x^2}{2}$. O resultado segue de tomar a exponencial em ambos os lados. □

²Na seção 5.2 provamos que σ é o push-forward da medida de Haar em $\mathcal{O}(d)$ e é a única medida invariante pelo grupo ortogonal $\mathcal{O}(d)$.

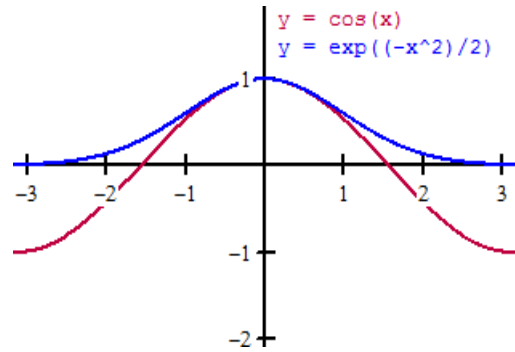


Figura 3.2: O gráfico representa a desigualdade $\cos(x) \leq \exp(-x^2/2)$.

Vamos calcular a área da calota $C = C_N^{d+1}(\epsilon) \subset S^{d+1} \subset \mathbb{R}^{d+2}$

$$\begin{aligned} \eta(C) &= \int_0^{2\pi} \cdots \int_0^\pi \int_0^{\pi/2-\theta} \text{sen}^d(\phi_1) \text{sen}^{d-1}(\phi_2) \cdots \text{sen}(\phi_d) d\phi_1 \cdots d\phi_{d+1} \\ &= \int_0^{2\pi} \cdots \int_0^\pi \text{sen}^{d-1}(\phi_2) \cdots \text{sen}(\phi_d) d\phi_2 \cdots d\phi_{d+1} \int_0^{\pi/2-\theta} \text{sen}^d(\phi_1) d\phi_1 \\ &= \eta(S^d) \int_0^{\pi/2-\theta} \text{sen}^d(\phi_1) d\phi_1. \end{aligned}$$

Como $\text{sen}(\phi_1) = \cos(\phi_1 - \frac{\pi}{2})$, fazendo $\phi = \phi_1 - \frac{\pi}{2}$ e $\frac{\Gamma(\frac{d+2}{2})}{\Gamma(\frac{d+1}{2})\sqrt{\pi}} = \alpha$, temos

$$\begin{aligned} \sigma(C) &= \frac{\eta(C)}{\eta(S^{d+1})} = \frac{\eta(S^d)}{\eta(S^{d+1})} \int_0^{\pi/2-\theta} \text{sen}^d(\phi_1) d\phi_1 \\ &= \alpha \int_0^{\pi/2-\theta} \text{sen}^d(\phi_1) d\phi_1 = \alpha \int_\theta^{\pi/2} \cos^d(\phi) d\phi. \end{aligned}$$

Usando $\phi = \frac{\psi}{\sqrt{d}}$ e aplicando a Proposição 3.4.1, temos

$$\sigma(C) \leq \frac{\alpha}{\sqrt{d}} \int_{\theta\sqrt{d}}^{\frac{\pi}{2}\sqrt{d}} e^{-\frac{\psi^2}{2}} d\psi.$$

Agora faça $\omega = \psi - \theta\sqrt{d}$

$$\begin{aligned} \sigma(C) &\leq \frac{\alpha}{\sqrt{d}} e^{-\frac{\theta^2 d}{2}} \int_0^{\sqrt{d}(\frac{\pi}{2}-\theta)} e^{-\frac{\omega^2}{2}} d\omega \\ &\leq \frac{\alpha}{\sqrt{d}} e^{-\frac{\theta^2 d}{2}} \int_0^\infty e^{-\frac{\omega^2}{2}} d\omega = \frac{\alpha}{\sqrt{d}} \frac{\sqrt{2\pi}}{2} e^{-\frac{\theta^2 d}{2}} \\ &= \frac{\Gamma(\frac{d+2}{2})}{\Gamma(\frac{d+1}{2})\sqrt{2d}} \exp\left(-\frac{\theta^2 d}{2}\right). \end{aligned}$$

Fazendo $s_d = \frac{\Gamma(\frac{d+2}{2})}{\Gamma(\frac{d+1}{2})\sqrt{2d}}$ e usando a equação $\Gamma(z+1) = z\Gamma(z)$, temos

$$\begin{aligned} s_{d+2} &= \frac{\Gamma(1 + \frac{d+2}{2})}{\Gamma(1 + \frac{d+1}{2})} \frac{1}{\sqrt{2(d+2)}} = \frac{\frac{d+2}{2}\Gamma(\frac{d+2}{2})}{\frac{d+1}{2}\Gamma(\frac{d+1}{2})} \frac{1}{\sqrt{2(d+2)}} \\ &= \left(\frac{\Gamma(\frac{d+2}{2})}{\Gamma(\frac{d+1}{2})\sqrt{2d}} \right) \frac{d+2}{d+1} \frac{\sqrt{2d}}{\sqrt{2(d+2)}} = s_d \frac{d+2}{d+1} \frac{\sqrt{2d}}{\sqrt{2(d+2)}} \leq s_d. \end{aligned}$$

Logo, calculando $\max\{s_1, s_2\} = \frac{2}{\sqrt{\pi}}$, temos $s_d \leq \frac{2}{\sqrt{\pi}}$, $\forall d \in \mathbb{N}$. Portanto,

$$\sigma(S^{d-1} \setminus E_\epsilon^{d-1}) = 2 \sigma(C) \leq \frac{4}{\sqrt{\pi}} \exp\left(-\frac{\theta^2 d}{2}\right).$$

Fazendo $d \rightarrow \infty$, provamos a concentração da esfera

$$\lim_{d \rightarrow \infty} \sigma(S^{d-1} \setminus E_\epsilon^{d-1}) = 0.$$

4

Análise em Grassmannianas

O teorema de Johnson-Lindenstrauss emprega as projeções de \mathbb{R}^d sobre subespaços vetoriais de dimensão k . O conjunto $G_k(\mathbb{R}^d)$ dos subespaços de dimensão k em \mathbb{R}^d , chamado de *Grassmanniana*, é o objeto central do capítulo.

Apresentamos $G_k(\mathbb{R}^d)$ como um espaço métrico compacto, exibindo-o depois como subvariedade do espaço das matrizes simétricas, apresentando cartas locais e um atlas diferenciável.

Neste capítulo, se $E \in G_k(\mathbb{R}^d)$ então P_E pode representar a projeção ortogonal sobre E ou a matriz de P_E na base canônica.

Seja

$$\mathcal{G}(d, k) := \{P \in \mathcal{M}(d, \mathbb{R}) : P^2 = P, P^* = P^T = P, \text{tr}(P) = k\},$$

onde tr é função traço. Pelo teorema espectral, toda matriz simétrica ($P^* = P$) e idempotente ($P^2 = P$) de posto k se decompõe como

$$P = Q\Lambda Q^*, \tag{4-1}$$

onde $Q \in \mathcal{O}(d)$ e Λ é a matriz diagonal cujas k primeiras entradas diagonais são iguais a 1, e as outras são nulas.

Como $\text{tr}(AB) = \text{tr}(BA)$, o traço é invariante por conjugação. Então uma matriz P idempotente e simétrica tem $\text{posto}(P) = k$ se, e somente se, $\text{tr}(P) = k$. Assim,

$$\mathcal{G}(d, k) = \{P \in \mathcal{M}(d, \mathbb{R}) : P^2 = P, P^* = P, \text{posto}(P) = k\}.$$

Proposição 4.0.2 *A função $h : G_k(\mathbb{R}^d) \rightarrow \mathcal{G}(d, k)$, $E \mapsto P_E$, é uma bijeção. O conjunto $\mathcal{G}(d, k)$ é um espaço métrico compacto.*

Assim, a identificação induz em $G_k(\mathbb{R}^d)$ uma estrutura de espaço métrico compacto. Na verdade, todas as propriedades de interesse da Grassmanniana serão estudadas em $\mathcal{G}(d, k)$, e meramente transferidas para a descrição em termos de subespaços.

Demonstração: Fixada a base canônica em \mathbb{R}^d , existe um isomorfismo entre $\mathcal{M}(d, \mathbb{R})$ e $\mathcal{L}(\mathbb{R}^d)$, o conjunto de todos os operadores lineares $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$. Lembramos que toda projeção ortogonal $P : \mathbb{R}^d \rightarrow \mathbb{R}^d$ sobre um subespaço $E \subset \mathbb{R}^d$ é autoadjunta, idempotente e satisfaz $\text{posto}(P) = \dim E$. Assim, para cada $E \in G_k(\mathbb{R}^d)$ designamos $P_E : \mathbb{R}^d \rightarrow \mathbb{R}^d$, a matriz(única) da projeção ortogonal sobre E , na base canônica. Logo, a função h , além de estar bem definida, é injetiva. Agora provaremos a sobrejetividade. Seja $P \in \mathcal{G}(d, k)$. Pela bijeção entre $\mathcal{M}(d, \mathbb{R})$ e $\mathcal{L}(\mathbb{R}^d)$, tomamos o operador $P_E : \mathbb{R}^d \rightarrow \mathbb{R}^d$ correspondente à matriz P que claramente é idempotente e autoadjunto. Além disso, $\text{Im}(P)$ tem dimensão k . Ou seja, P é uma projeção ortogonal sobre $E := \text{Im}P \in G_k(\mathbb{R}^d)$. Logo $h(E) = P$.

A norma de Frobenius de $A \in \mathcal{M}(d, \mathbb{R})$ é dada por $\|A\|^2 = \text{tr}(A^*A)$. Para $P \in \mathcal{G}(d, k)$

$$\|P\|^2 = \text{tr}(P_E^*P_E) = \text{tr}(P_E^2) = \text{tr}(P_E) = k,$$

mostrando que $\mathcal{G}(d, k)$ é limitado.

Para ver que $\mathcal{G}(d, k)$ é fechado, note que ele é a interseção de níveis de várias funções contínuas (os níveis zero de $(P^2 - P)$, de $(P^* - P)$ e de $(\text{tr}P - k)$).

Portanto, $\mathcal{G}(d, k)$ é fechado e mais, compacto. \square

4.1

$G_k(\mathbb{R}^d)$ como subvariedade de $\mathcal{S}(\mathbb{R}^d)$

Considere a ação por conjugação do grupo $\mathcal{O}(d)$ sobre o conjunto das matrizes reais simétricas $\mathcal{S}(\mathbb{R}^d)$

$$\begin{aligned} \phi : \mathcal{O}(d) \times \mathcal{S}(\mathbb{R}^d) &\rightarrow \mathcal{S}(\mathbb{R}^d) \\ (Q, S) &\mapsto QSQ^*. \end{aligned}$$

Se duas matrizes simétricas S e Λ têm os mesmos autovalores então, pelo teorema espectral, as orbitas da ação por conjugação passando por S e Λ são iguais. Para a matriz Λ descrita em (4-1), as matrizes da órbita são todas idempotentes, simétricas e do mesmo traço. Assim a órbita de Λ coincide com $\mathcal{G}(d, k)$.

Usaremos a Proposição 2.0.1, anunciada no capítulo 2, para provar que $\mathcal{G}(d, k) = \mathcal{O}(d)\Lambda$ é uma subvariedade fechada de $\mathcal{S}(\mathbb{R}^d)$:

Proposição 1 *Seja $\phi : G \times M \rightarrow M$ uma ação própria. Então toda órbita Gx é uma subvariedade fechada de M .*

Lembramos ao leitor que uma ação ϕ é própria se $\tilde{\phi} : G \times M \rightarrow M \times M$, $(g, x) \mapsto (x, \phi(g, x))$, é uma aplicação própria, isto é, se a inversa de um compacto é compacto.

Corolário 4.1 $\mathcal{G}(d, k)$ é uma subvariedade fechada de $\mathcal{S}(\mathbb{R}^d)$.

Demonstração: Neste caso, $\tilde{\phi} : \mathcal{O}(d) \times \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d) \times \mathcal{S}(\mathbb{R}^d)$ é dada por $\tilde{\phi}(Q, S) = (S, \phi(Q, S)) = (S, QSQ^*)$. Tome $K = K_1 \times K_2 \subset \mathcal{S}(\mathbb{R}^d) \times \mathcal{S}(\mathbb{R}^d)$ com K_1 e K_2 compactos. Provaremos que

$$\tilde{K} = (\tilde{\phi})^{-1}(K) = \{(Q, S) \in \mathcal{O}(d) \times \mathcal{S}(\mathbb{R}^d) : S \in K_1, \phi(Q, S) \in K_2\}$$

é um compacto. É claro que \tilde{K} é fechado, pois $\tilde{\phi}$ é contínua e K é compacto. Logo, resta mostrar que \tilde{K} é limitado. Ora, $\tilde{K} \subset \mathcal{O}(d) \times K_1$ onde $\mathcal{O}(d)$ e K_1 , ambos limitados. Logo, \tilde{K} é limitado. \square

4.2

A função posto

Seja uma $A = (a_{ij})$ uma matriz de dimensão $m \times n$. Nesta seção, A_{ij} é a submatriz de A de dimensão $i \times j$ com entradas na interseção das i primeiras linhas com as j primeiras colunas de A . O exemplo é esclarecedor.

$$A = \left(\begin{array}{cc|cc} 3 & 1 & 1 & 0 \\ 2 & 1 & 2 & 8 \\ \hline 2 & 2 & 4 & 7 \\ 6 & 5 & 1 & 0 \\ 3 & 1 & 4 & 2 \end{array} \right)_{5 \times 4} \quad A_{32} = \left(\begin{array}{cc} 3 & 1 \\ 2 & 1 \\ 2 & 2 \end{array} \right)_{3 \times 2} .$$

Seja $A = (a_{ij})$ uma matriz $m \times n$. Definimos a *função posto* \mathcal{P}_A

$$\begin{aligned} \mathcal{P}_A : \{1, 2, \dots, m\} \times \{1, 2, \dots, n\} &\rightarrow \mathbb{N} \cup \{0\} \\ (i, j) &\mapsto \text{posto}(A_{ij}) \end{aligned}$$

Damos uma visualização matricial à função \mathcal{P}_A . A *matriz posto* de A é a matriz $P_A = (p_{ij})_{m \times n}$ onde $p_{ij} = \mathcal{P}_A(i, j)$.

Exemplo:

$$A = \left(\begin{array}{ccc} 0 & 0 & 3 \\ 2 & 1 & 2 \\ 2 & 1 & 4 \\ 6 & 5 & 1 \end{array} \right)_{4 \times 3} \quad P_A = \left(\begin{array}{ccc} 0 & 0 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 2 & 3 \end{array} \right)_{4 \times 3} .$$

A configuração local em ij da matriz P_A é a submatriz

$$\begin{array}{c|c} p_{(i-1)(j-1)} & p_{(i-1)j} \\ \hline p_{i(j-1)} & p_{ij} \end{array}.$$

Para a matriz A acima, a configuração local em 23 é $\begin{array}{c|c} 0 & 1 \\ \hline 1 & 2 \end{array}$.

Observamos duas propriedades simples da função posto: na primeira linha ou coluna da matriz A a função \mathcal{P}_A toma valores iguais a 0 ou 1 e quando a matriz aumenta uma linha ou uma coluna, o posto pode aumentar em no máximo 1. Ou seja, \mathcal{P}_A é monótona, no seguinte sentido

$$p_{ij} \leq p_{(i+1)j} \leq p_{ij} + 1, \quad p_{ij} \leq p_{i(j+1)} \leq p_{ij} + 1.$$

Os seguintes diagramas mostram as configurações locais possíveis:

$$\begin{array}{c|c} p & p \\ \hline p & p \end{array}, \quad \begin{array}{c|c} p & p \\ \hline p & p+1 \end{array}, \quad \begin{array}{c|c} p & p+1 \\ \hline p & p+1 \end{array},$$

$$\begin{array}{c|c} p & p \\ \hline p+1 & p+1 \end{array}, \quad \begin{array}{c|c} p & p+1 \\ \hline p+1 & p+1 \end{array}, \quad \begin{array}{c|c} p & p+1 \\ \hline p+1 & p+2 \end{array}.$$

Porém a proposição seguinte descarta a quinta configuração.

Proposição 4.2.1 *Seja $A = (a_{ij})_{m \times n}$ uma matriz e $P_A = (p_{ij})_{m \times n}$ sua matriz posto. Se $p_{ij} = p$, $p_{i(j+1)} = p + 1$ e $p_{(i+1)j} = p + 1$, então $p_{(i+1)(j+1)} = p + 2$.*

Demonstração: Ao longo desta prova, para cada $v \in \mathbb{R}^{i+1}$, $v' \in \mathbb{R}^i$ denotará o vetor obtido truncando a última coordenada de v . A demonstração é por absurdo. Como

$$P_{A_{(i+1)(j+1)}} = \left(\begin{array}{c|c} & \\ \hline & p & p+1 \\ \hline p+1 & p+1 \end{array} \right) = \left(\begin{array}{c|c} P_{A_{(i+1)j}} & * \\ \hline p+1 & * \end{array} \right),$$

existem $p + 1$ colunas LI em $A_{(i+1)j}$.

Sejam $\{v_1, v_2, \dots, v_p, v_{p+1}\}$ tais vetores em \mathbb{R}^{i+1} e $\{v'_1, v'_2, \dots, v'_p, v'_{p+1}\}$ seus truncamentos em \mathbb{R}^i .

Fato 1. *Exatamente p vetores de $\{v'_1, v'_2, \dots, v'_p, v'_{p+1}\}$ são LI.*

Demonstração: Seja $k < p$ e suponha que no máximo k vetores de $\{v'_1, v'_2, \dots, v'_p, v'_{p+1}\}$ sejam LI. Sem perda de generalidade sejam

$\{v'_1, v'_2, \dots, v'_k\}$ tais vetores. Temos então que

$$\dim(\text{span}\{v'_1, v'_2, \dots, v'_p, v'_{p+1}\}) = k.$$

Logo

$$\dim(\text{span}\{v_1, v_2, \dots, v_p, v_{p+1}\}) \leq k + 1 \leq p,$$

um absurdo pois $\{v_1, v_2, \dots, v_p, v_{p+1}\}$ é um conjunto LI. Por outro lado, $\{v'_1, v'_2, \dots, v'_p, v'_{p+1}\}$ não pode ser LI, pois $\{v'_1, v'_2, \dots, v'_p, v'_{p+1}\}$ são colunas da submatriz A_{ij} que tem posto p . Assim concluímos que exatamente p vetores de $\{v'_1, v'_2, \dots, v'_p, v'_{p+1}\}$ são LI. \lrcorner

Sem perda, sejam $\{v'_1, v'_2, \dots, v'_p\}$ esses p vetores. Então

$$v'_{p+1} = \beta_1 v'_1 + \beta_2 v'_2 + \dots + \beta_p v'_p. \quad (4-2)$$

Seja w a última coluna da submatriz $A_{(i+1)(j+1)}$ e considere seu truncamento w' , a última coluna de $A_{i(j+1)}$

$$P_{A_{(i+1)(j+1)}} = \left(\begin{array}{ccc} & & | \\ & & | \\ & P_{A_{i(j+1)}} & | \ P_{w'} \\ & & | \\ & & | \\ \hline & & | \ p+1 \end{array} \right).$$

Fato 2. $\{v'_1, v'_2, \dots, v'_p, w'\}$ é um conjunto LI.

Demonstração: Por hipótese $p_{i(j+1)} = p+1$. Logo a submatriz $A_{i(j+1)}$ tem $p+1$ colunas LI, onde necessariamente, w' é uma dessas colunas. Caso contrário A_{ij} teria $p+1$ colunas LI, contradizendo o fato da submatriz ter posto p . Tomamos então essas $p+1$ colunas LI como sendo $\{u_1, \dots, u_p, w'\}$. Ora, $\{u_1, \dots, u_p\}$ e $\{v'_1, v'_2, \dots, v'_p\}$ são duas bases para o subespaço gerado pelas colunas de A_{ij} , onde w' não pertence a esse subespaço, pois $\{w', u_1, \dots, u_p\}$ é um conjunto LI. Assim, $\{v'_1, v'_2, \dots, v'_p, w'\}$ é LI. \lrcorner

Como $\{v_1, v_2, \dots, v_p, v_{p+1}\}$ é LI em \mathbb{R}^{i+1} , temos, para α' s adequados,

$$w = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_p v_p + \alpha_{p+1} v_{p+1}$$

e, em \mathbb{R}^i ,

$$w' = \alpha_1 v'_1 + \alpha_2 v'_2 + \dots + \alpha_p v'_p + \alpha_{p+1} v'_{p+1}. \quad (4-3)$$

Substituindo (4-2) em (4-3),

$$\begin{aligned} w' &= \alpha_1 v'_1 + \alpha_2 v'_2, \dots, \alpha_p v'_p + \alpha_{p+1} (\beta_1 v'_1 + \beta_2 v'_2, \dots, \beta_p v'_p) \\ &= (\alpha_1 + \alpha_{p+1} \beta_1) v'_1 + \dots + (\alpha_p + \alpha_{p+1} \beta_p) v'_p, \end{aligned}$$

um absurdo, pois $\{v'_1, v'_2, \dots, v'_p, w'\}$ é LI, pelo fato anterior. \square

Exemplo: A existência de cada uma das configurações locais remanescentes

$$\frac{p}{p} \left| \frac{p}{p} \right.; \quad \frac{p}{p} \left| \frac{p}{p+1} \right.; \quad \frac{p}{p} \left| \frac{p+1}{p+1} \right.; \quad \frac{p}{p+1} \left| \frac{p}{p+1} \right.; \quad \frac{p}{p+1} \left| \frac{p+1}{p+2} \right.$$

pode ser verificada a seguir

$$A = \begin{pmatrix} 0 & 0 & 1 & 9 & 0 & 1 \\ 3 & 8 & 6 & 21 & 16 & \mathbf{7} \\ 0 & 5 & \mathbf{2} & 7 & 10 & 53 \\ 7 & 1 & 32 & 0 & 2 & 4 \\ 0 & 0 & 2 & 18 & \mathbf{0} & 2 \\ 2 & 3 & 7 & 8 & 6 & 1 \\ 0 & \mathbf{3} & 84 & 2 & 6 & \mathbf{0} \end{pmatrix}_{7 \times 6} \quad P_A = \begin{pmatrix} 0 & 0 & 1 & 1 & \mathbf{1} & \mathbf{1} \\ 1 & \mathbf{1} & \mathbf{2} & 2 & \mathbf{2} & \mathbf{2} \\ 1 & \mathbf{2} & \mathbf{3} & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 & 4 \\ 1 & 2 & 3 & 4 & 4 & 4 \\ \mathbf{1} & \mathbf{2} & 3 & 4 & 4 & 4 \\ \mathbf{1} & \mathbf{2} & 3 & 4 & 4 & \mathbf{5} \end{pmatrix}_{7 \times 6}.$$

Para 66, 62, 55, 33, 26 as configurações locais são

$$\frac{4}{4} \left| \frac{4}{5} \right.; \quad \frac{1}{1} \left| \frac{2}{2} \right.; \quad \frac{4}{4} \left| \frac{4}{4} \right.; \quad \frac{1}{2} \left| \frac{2}{3} \right.; \quad \frac{1}{2} \left| \frac{1}{2} \right.$$

Vamos emoldurar A com uma primeira linha e uma primeira coluna de zeros, como abaixo. As entradas da *matriz aumentada* $\tilde{A} = (\tilde{a}_{ij})_{(m+1) \times (n+1)}$ são indexadas a partir de 0

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Seja $A = (a_{ij})$ uma matriz $m \times n$ e $\tilde{A} = (\tilde{a}_{ij})$ sua matriz aumentada. Vamos *marcar* algumas entradas de A : as entradas ij de \tilde{A} cuja configuração

local é dada por $\frac{p}{p} \left| \frac{p}{p+1} \right.$.

A *configuração de torres* de A é a matriz $\mathcal{T}(A)$ de zeros e uns, da

mesma dimensão de A , onde os uns indicam as entradas marcadas em A . A palavra *torre* é empregada para sugerir que, como veremos no primeiro item da Proposição 4.2.2, não é possível ter duas entradas marcadas ou na mesma linha ou na mesma coluna, isto é, duas torres não se atacam.

Exemplo: Seja A a matriz do exemplo anterior. A matriz aumentada \tilde{A} , a matriz posto de \tilde{A} e a configuração de torres de A são, respectivamente,

$$\tilde{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 9 & 0 & 1 \\ 0 & 3 & 8 & 6 & 21 & 16 & 7 \\ 0 & 0 & 5 & 2 & 7 & 10 & 53 \\ 0 & 7 & 1 & 32 & 0 & 2 & 4 \\ 0 & 0 & 0 & 2 & 18 & 0 & 2 \\ 0 & 2 & 3 & 7 & 8 & 6 & 1 \\ 0 & 0 & 3 & 84 & 2 & 6 & 0 \end{pmatrix}_{8 \times 7}$$

$$P_{\tilde{A}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 1 & 1 & 1 \\ 0 & \mathbf{1} & 1 & 2 & 2 & 2 & 2 \\ 0 & 1 & \mathbf{2} & 3 & 3 & 3 & 3 \\ 0 & 1 & 2 & 3 & \mathbf{4} & 4 & 4 \\ 0 & 1 & 2 & 3 & 4 & \mathbf{4} & 4 \\ 0 & 1 & 2 & 3 & 4 & 4 & \mathbf{4} \\ 0 & 1 & 2 & 3 & 4 & 4 & \mathbf{5} \end{pmatrix}_{8 \times 7}$$

$$\mathcal{T}(A) = \begin{pmatrix} 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} \end{pmatrix}_{7 \times 6}.$$

É claro que, se na primeira linha (ou coluna) de qualquer matriz não existem torres, então as entradas dessa linha(ou coluna) são todas iguais a 0.

Lema 4.2 *Seja $A = (a_{ij})$ uma matriz $m \times n$. Suponha que na linha $r > 1$ não existam torres. Então, em $P_A = (p_{ij})$ temos $p_{rj} = p_{(r-1)j}$, $\forall j = 1, 2, \dots, n$. Ou seja, a linha r da matriz P_A é uma cópia da sua antecessora. O resultado é análogo para colunas.*

Demonstração: Mostramos o resultado para linhas. Tome a primeira coluna $k \in \{1, 2, \dots, n\}$ com $p_{rk} \neq p_{(r-1)k}$. Se $p_{(r-1)k} = p$, então pela monotonicidade da função posto, a configuração local na entrada rk é

$$\begin{array}{c|c} p_{(r-1)(k-1)} & p \\ \hline p_{r(k-1)} & \mathbf{p+1} \end{array}$$

e os únicos valores possíveis para $p_{(r-1)(k-1)}$ e $p_{r(k-1)}$ são

$$\frac{p}{p} \mid \frac{p}{\mathbf{p}+1},$$

o que implicaria a existência de uma torre na linha r , um absurdo. \square

Proposição 4.2.2 *Valem as seguintes propriedades:*

1. *As torres de uma matriz $A = (a_{ij})_{m \times n}$ não se atacam. Ou seja, em cada linha ou coluna da matriz A , existe no máximo uma torre.*
2. *O número de torres de qualquer matriz $A = (a_{ij})_{m \times n}$ é igual ao posto da mesma.*

Demonstração: Provaremos só o caso das linhas, o outro é análogo. Suponha que na entrada a_{rk} exista uma torre

$$\frac{a_{(r-1)(k-1)} \mid a_{(r-1)k}}{a_{r(k-1)} \mid \mathbf{a}_{rk}} \quad \frac{p}{p} \mid \frac{p}{\mathbf{p}+1}.$$

Dentro das possíveis configurações, existem só duas possibilidades ao preencher a coluna $k + 1$ nas linhas r e $r - 1$

$$\frac{p}{p+1} \mid \frac{p}{p+1}, \quad \frac{p}{p+1} \mid \frac{p+1}{p+2}.$$

Repetindo este processo, concluímos que não pode haver outra torre na mesma linha r . Assim, provamos o primeiro item. A prova do segundo item será feita por indução na variável m fixando a variável n . Para $m = 1$ o resultado é bastante claro. Suponha que para toda matriz de dimensão $m \times n$ o número de torres é igual ao posto. Seja então uma matriz $A = (a_{ij})_{(m+1) \times n}$. Nosso objetivo é provar que A tem posto igual ao número de torres. Considere a matriz posto de A

$$P_A = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mn} \\ p_{(m+1)1} & p_{(m+1)2} & \cdots & p_{(m+1)n} \end{pmatrix}.$$

Se retirarmos a última linha da matriz A obtemos uma submatriz A_{mn} . Pela hipótese de indução o número de torres de A_{mn} é igual ao seu

posto. Assumimos $p_{(m+1) \times n} \neq 0$. Caso contrário o resultado é trivial. Assim, denotamos $p_{(m+1) \times n} = p + 1$

$$P_A = \left(\begin{array}{c|c} p_{mn} \text{ torres} & \\ \hline & \begin{array}{c} p_{mn} \\ p + 1 \end{array} \end{array} \right)_{(m+1) \times n} .$$

Pela monotonicidade da função posto, temos dois possíveis valores para p_{mn} . A primeira é $p_{mn} = p$

$$P_A = \left(\begin{array}{c|c} p \text{ torres} & \\ \hline & \begin{array}{c} p \\ p + 1 \end{array} \end{array} \right)_{(m+1) \times n} .$$

Neste caso, necessariamente existe uma torre na ultima linha, caso contrário, pelo Lema 4.2 teriamos $p_{mn} = p_{(m-1)n}$. Assim, o numero de torres da matriz A é igual ao seu posto. Finalmente, a outra possibilidade é $p_{mn} = p + 1$

$$P_A = \left(\begin{array}{c|c} p + 1 \text{ torres} & \\ \hline & \begin{array}{c} p + 1 \\ p + 1 \end{array} \end{array} \right)_{(m+1) \times n} .$$

Neste caso, não pode existir uma torre na ultima linha P_A . Se existe uma torre nessa linha, usamos o mesmo argumento usado na prova do primeiro item para concluir que $p_{(m+1)n} = p_{mn} + 1$, um absurdo. \square

Corolário 4.3 *Toda matriz A quadrada de dimensão n é inversível se, e somente se, A tem n torres. Além disso, ao permutarmos duas linhas ou duas colunas de uma matriz qualquer, então o número de torres não se altera.*

Demonstração: Este resultado é consequência imediata do segundo item da proposição anterior. Se uma matriz quadrada A de dimensão n é inversível, então $\text{posto}(A) = n$. Assim, A tem n torres. Reciprocamente, se $A = (a_{ij})$ é uma matriz quadrada de dimensão n possuindo n torres então $\text{posto}(A) = n$, logo A é inversível. Ao permutarmos duas linhas ou colunas da matriz A o seu posto não muda, logo o número de torres se mantém. \square

Proposição 4.2.3 *Seja Q uma matriz $d \times k$ com $d \geq k$. Se Q possui k torres, ou seja, uma torre em cada coluna, então existe uma matriz permutação Π de*

modo que ΠQ tem a seguinte configuração de torres.

$$\mathcal{T}(\Pi Q) = \left(\begin{array}{cccc|c} \mathbf{1} & 0 & 0 & 0 & \\ 0 & \mathbf{1} & 0 & 0 & \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \mathbf{1} & \\ \hline & & & & 0 \end{array} \right)_{d \times k} .$$

Demonstração: Para qualquer permutação, ou sequência de permutações, das linhas de Q , corresponde uma matriz permutação. Lembramos a notação de submatriz: Q_{dj} representa a submatriz de Q cujas colunas são exatamente as j primeiras colunas de Q .

Por hipótese, na primeira coluna de Q existe uma torre. Logo, nessa mesma coluna, existe uma entrada diferente de zero. Suponha que essa entrada não nula se encontre na linha $r \neq 1$ da matriz Q . Ao permutarmos a linha r com a primeira linha de Q obtemos uma matriz permutação Π_1 tal que $Q^1 = \Pi_1 Q = (q_{ij}^1)$ tem uma torre na entrada q_{11}^1 . Além disso, Q^1 continua com k torres e uma em cada coluna. Agora tomemos a submatriz Q_{d2}^1 . Sabemos que existem duas torres na submatriz Q_{d2}^1 onde uma delas se encontra na posição $(1, 1)$. Suponha que a segunda torre de Q_{d2}^1 se encontra na posição $(r, 2)$. Ora, pelo segundo item do Proposição 4.2.2, a submatriz Q_{d2}^1 tem posto 2, isto é, Q_{d2}^1 tem duas linhas linearmente independentes. Pelo Lema 4.2, é fácil concluir que todas as linhas de Q_{d2}^1 entre as duas torres são linearmente dependentes da primeira. Concluimos assim que as linhas 1 e r da submatriz Q_{d2}^1 são LI e ao permutarmos as linhas 2 e r , obtemos uma matriz permutação Π_2 tal que $Q^2 = \Pi_2 Q^1 = \Pi_2 \Pi_1 Q = (q_{ij}^2)$ tem duas torres nas entradas q_{11}^2 e q_{22}^2 . Além disso, Q^2 continua com k torres e uma em cada coluna. Continuando este processo, obtemos uma matriz permutação $\Pi = \Pi_k \circ \Pi_{k-1} \circ \dots \circ \Pi_1$ tal que $\Pi Q = Q^k = (q_{ij}^k)$ tem as k torres em $q_{11}^k, q_{22}^k, \dots, q_{kk}^k$. \square

4.3

$G_k(\mathbb{R}^d)$ como variedade diferenciável

Já provamos que $G_k(\mathbb{R}^d)$ é uma subvariedade de $\mathcal{S}(\mathbb{R}^d)$. Nesta seção exibiremos cartas.

Para $Q \in \mathcal{O}(d)$, defina Q_k a matriz com as k primeiras colunas de Q . Se A é uma matriz quadrada de dimensão k , escrevemos $\mathcal{T}(A) = \mathbb{I}_k$, para indicar

que A tem as torres na diagonal, ou seja, a configuração de torres de A é igual à configuração de torres da matriz identidade de dimensão k .

Se $E \in G_k(\mathbb{R}^d)$ e P_E é a matriz da projeção ortogonal sobre E , lembramos ao leitor que pelo teorema espectral

$$P_E = Q\Lambda Q^*,$$

onde $Q \in \mathcal{O}(d)$ e Λ é a matriz diagonal cujas k primeiras entradas diagonais são iguais a 1, e as outras são nulas. As colunas de Q_k formam uma base de E , logo $\text{posto}(Q_k) = k$. Portanto, Q_k tem k torres e, pela Proposição 4.2.3, existe uma matriz de permutação Π tal que

$$\Pi Q_k = \begin{bmatrix} A \\ B \end{bmatrix}, \quad \mathcal{T}(A) = \mathbb{I}_k.$$

Finalmente, pelo Corolário 4.3, A é inversível. Feito este procedimento, determinamos as cartas locais.

Sejam $\mathcal{I}_d = \{1, 2, \dots, d\}$ e S_d o grupo simétrico correspondente, cujos elementos são as permutações de \mathcal{I}_d . Para cada $\pi \in S_d$, denota-se Π_π a matriz de permutação correspondente. Seja

$$U_\pi = \{E \in G_k(\mathbb{R}^d) : P_E = Q\Lambda Q^*, \Pi_\pi Q_k = \begin{bmatrix} A \\ B \end{bmatrix}, \mathcal{T}(A) = \mathbb{I}_k\}.$$

$\mathcal{M}((d-k) \times k)$ é o conjunto das matrizes $(d-k) \times k$. As cartas são definidas como

$$\begin{aligned} \Phi_\pi : U_\pi \subset G_k(\mathbb{R}^d) &\rightarrow \mathcal{M}((d-k) \times k) \\ E &\mapsto BA^{-1}. \end{aligned}$$

A matriz Q , na decomposição espectral de P_E , não é única.

Proposição 4.3.1 *A função Φ_π é bem definida.*

Demonstração: Sejam $E \in U_\pi$ e P_E a projeção ortogonal sobre E . Considere duas decomposições de P_E , ou seja, $P_E = Q\Lambda Q^* = \tilde{Q}\tilde{\Lambda}\tilde{Q}^*$ com $\Pi_\pi Q_k = \begin{bmatrix} A \\ B \end{bmatrix}$, $\Pi_\pi \tilde{Q}_k = \begin{bmatrix} \tilde{A} \\ \tilde{B} \end{bmatrix}$ e $\mathcal{T}(A) = \mathcal{T}(\tilde{A}) = \mathbb{I}_k$. Se na igualdade $Q\Lambda Q^* = \tilde{Q}\tilde{\Lambda}\tilde{Q}^*$ multiplicamos \tilde{Q}^* pela esquerda e Q pela direita, temos

$$W\Lambda = \Lambda W, \quad W = \tilde{Q}^*Q \in \mathcal{O}(d). \quad (4-4)$$

Se $W = \left[\begin{array}{c|c} W_{11} & W_{12} \\ \hline W_{21} & W_{22} \end{array} \right]_{d \times d}$, com $W_{11} \in \mathcal{M}(k)$, então

$$\begin{aligned} W\Lambda &= \Lambda W \\ \left[\begin{array}{c|c} W_{11} & W_{12} \\ \hline W_{21} & W_{22} \end{array} \right] \left[\begin{array}{c|c} I_{k \times k} & 0 \\ \hline 0 & 0 \end{array} \right] &= \left[\begin{array}{c|c} I_{k \times k} & 0 \\ \hline 0 & 0 \end{array} \right] \left[\begin{array}{c|c} W_{11} & W_{12} \\ \hline W_{21} & W_{22} \end{array} \right] \\ \left[\begin{array}{c|c} W_{11} & 0 \\ \hline W_{21} & 0 \end{array} \right] &= \left[\begin{array}{c|c} W_{11} & W_{12} \\ \hline 0 & 0 \end{array} \right]. \end{aligned}$$

Logo $W_{12} = W_{21} = 0$. Assim

$$W = \left[\begin{array}{c|c} W_{11} & 0 \\ \hline 0 & W_{22} \end{array} \right]. \quad (4-5)$$

Multiplicando \tilde{Q} pela esquerda na igualdade $W = \tilde{Q}^*Q$, temos

$$Q = \tilde{Q}W. \quad (4-6)$$

Se $\tilde{Q} = \left[\begin{array}{c|c} \tilde{Q}_{11} & \tilde{Q}_{12} \\ \hline \tilde{Q}_{21} & \tilde{Q}_{22} \end{array} \right]$, então $\tilde{Q}W = \left[\begin{array}{c|c} \tilde{Q}_{11}W_{11} & \tilde{Q}_{12}W_{22} \\ \hline \tilde{Q}_{21}W_{11} & \tilde{Q}_{22}W_{22} \end{array} \right]$. Pondo

$Q = \left[\begin{array}{c|c} Q_{11} & Q_{12} \\ \hline Q_{21} & Q_{22} \end{array} \right]$, com $Q_{11} \in \mathcal{M}(k)$, a equação (4-6) implica

$$\left[\begin{array}{c|c} Q_{11} & Q_{12} \\ \hline Q_{21} & Q_{22} \end{array} \right] = \left[\begin{array}{c|c} \tilde{Q}_{11}W_{11} & \tilde{Q}_{12}W_{22} \\ \hline \tilde{Q}_{21}W_{11} & \tilde{Q}_{22}W_{22} \end{array} \right].$$

Portanto

$$Q_k = \tilde{Q}_k W_{11}. \quad (4-7)$$

Multiplicando Π_π pela direita, temos

$$\begin{aligned} \Pi_\pi Q_k &= \Pi_\pi \tilde{Q}_k W_{11} \\ \left[\begin{array}{c} A \\ B \end{array} \right] &= \left[\begin{array}{c} \tilde{A}W_{11} \\ \tilde{B}W_{11} \end{array} \right]. \end{aligned}$$

Isto é, $A = \tilde{A}W_{11}$ e $B = \tilde{B}W_{11}$. Como W é ortogonal, da igualdade (4-5),

concluimos que W_{11} também é ortogonal e consequentemente inversível. Assim,

$$B A^{-1} = \tilde{B}W_{11}(\tilde{A}W_{11})^{-1} = \tilde{B}W_{11}W_{11}^{-1}\tilde{A}^{-1} = \tilde{B}\tilde{A}^{-1}.$$

Portanto, Φ_π é bem definida. O seguinte diagrama revisa o procedimento feito

$$\begin{array}{ccccc}
 & & Q\Lambda Q^* & \longrightarrow & \Pi_\sigma Q_k = \begin{bmatrix} A \\ B \end{bmatrix} \\
 & \nearrow & & & \searrow \\
 E \in G_k(\mathbb{R}^d) & \longrightarrow & P_E & & B A^{-1} = \tilde{B}\tilde{A}^{-1} \\
 & \searrow & & & \nearrow \\
 & & \tilde{Q}\Lambda\tilde{Q}^* & \longrightarrow & \Pi_\sigma \tilde{Q}_k = \begin{bmatrix} \tilde{A} \\ \tilde{B} \end{bmatrix}
 \end{array}$$

□

Agora mostraremos que Φ_π é bijetiva. Começamos com a injetividade. Sejam P_E e P_F , respectivamente, duas projeções ortogonais sobre E e $F \in U_\pi$. Suponha que $\Phi_\pi(E) = \Phi_\pi(F)$

$$\begin{array}{ccccc}
 & & E & \longrightarrow & P_E = Q\Lambda Q^* & \longrightarrow & \Pi_\pi Q_k = \begin{bmatrix} A \\ B \end{bmatrix} \\
 & \nearrow & & & & & \searrow \\
 E, F \in U_\pi & & & & & & B A^{-1} = \tilde{B}\tilde{A}^{-1} \\
 & \searrow & & & & & \nearrow \\
 & & F & \longrightarrow & P_F = \tilde{Q}\Lambda\tilde{Q}^* & \longrightarrow & \Pi_\pi \tilde{Q}_k = \begin{bmatrix} \tilde{A} \\ \tilde{B} \end{bmatrix}
 \end{array}$$

Tome $S = \tilde{A}^{-1}A$, então da igualdade $\tilde{B}\tilde{A}^{-1} = B A^{-1}$, temos $\tilde{B}S = \tilde{B}(\tilde{A}^{-1}A) = (\tilde{B}\tilde{A}^{-1})A = (B A^{-1})A = B$. Além disso, $\tilde{A}S = A$. Logo

$$\Pi_\pi Q_k = \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \tilde{A}S \\ \tilde{B}S \end{bmatrix} = \begin{bmatrix} \tilde{A} \\ \tilde{B} \end{bmatrix} S = \Pi_\pi \tilde{Q}_k S.$$

Portanto, $Q_k = \tilde{Q}_k S$ – equivalentemente $Q_k S^{-1} = \tilde{Q}_k$ –. Ou seja, cada coluna de \tilde{Q}_k é combinação linear das colunas de Q_k . Lembramos que as colunas de Q_k e \tilde{Q}_k são, respectivamente, bases para E e F . Logo, $E = F$. Assim, Φ_π é injetiva.

Agora mostraremos que Φ_π é sobrejetora. Seja $Z \in \mathcal{M}((d - k) \times k)$.

Tomamos I_k , a matriz identidade $k \times k$, e fazemos $M = \begin{bmatrix} I_k \\ Z \end{bmatrix}_{d \times k}$. Lembramos ao leitor que o processo de ortogonalização de Gram-Schmidt, aplicado às colunas de M , nos fornece uma matriz triangular superior $R_{k \times k}$, cujas entradas na diagonal são todas diferentes de 0 (logo inversível), e uma matriz ortogonal $\bar{Q}_k = \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix}_{d \times k}$, tal que

$$\begin{bmatrix} I_k \\ Z \end{bmatrix} = \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix} [R] = \begin{bmatrix} \bar{A}R \\ \bar{B}R \end{bmatrix}. \quad (4-8)$$

Ora, da igualdade $\bar{A}R = I_k$ segue que $\bar{A} = R^{-1}$. Dado que R^{-1} também é triangular superior, com todas as entradas da diagonal não nulas, concluímos que $\mathcal{T}(\bar{A}) = \mathbb{I}_k$.

Se $Q_k = \Pi_\pi^{-1} \bar{Q}_k$, então Q_k é ortogonal, pois é produto de duas matrizes ortogonais. Tomamos, então, E como sendo o subespaço vetorial gerado pelas colunas de Q_k . Desta maneira, temos $P_E = Q \Lambda Q^*$ com $Q = Q_k Q_k^* \in \mathcal{O}(d)$. Afirmamos que $\Phi_J(E) = Z$

$$E \mapsto P_E = Q \Lambda Q^* \mapsto \Pi_\pi Q_k = \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix} \mapsto \bar{B} \bar{A}^{-1}.$$

Usando (4-8), decretamos a sobrejetividade de Φ_π

$$\bar{B} \bar{A}^{-1} = \bar{B} R (R^{-1} \bar{A}^{-1}) = \bar{B} R (\bar{A} R)^{-1} = \bar{B} R I_k = \bar{B} R = Z.$$

□

Finalmente, mostramos a diferenciabilidade das mudanças de coordenadas.

Proposição 4.3.2 *Sejam $\Phi_\pi : U_\pi \rightarrow \mathcal{M}((d-k) \times k)$ e $\Phi_\mu : U_\mu \rightarrow \mathcal{M}((d-k) \times k)$ duas cartas locais. Então a mudança de coordenadas*

$$\begin{aligned} \Phi_\mu \circ \Phi_\pi^{-1} : \Phi_\pi(U_\pi \cap U_\mu) &\rightarrow \Phi_\mu(U_\pi \cap U_\mu) \\ Z &\mapsto \Phi_\mu(\Phi_\pi^{-1}(Z)) \end{aligned}$$

é suave.

Demonstração: Denotemos $\Phi_{\pi\mu} = \Phi_\mu \circ \Phi_\pi^{-1}$. Sejam $Z \in \Phi_\pi(U_\pi \cap U_\mu)$ e $E = \Phi_\pi^{-1}(Z) \in U_\pi \cap U_\mu$. Seguindo exatamente o mesmo argumento feito para a matriz Z em (4-8), tomamos $\bar{Q}_k = \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix}$ e $Q_k = \Pi_\pi^{-1}\bar{Q}_k$. Dado que $E \in U_\pi$, temos $P_E = Q\Lambda Q^*$ tal que $\Pi_\pi Q_k = \begin{bmatrix} A \\ B \end{bmatrix}$, $\mathcal{T}(A) = \mathbb{I}_k$. Por outro lado, se $E \in U_\mu$, temos $P_E = \tilde{Q}\Lambda\tilde{Q}^*$ tal que $\Pi_\mu\tilde{Q}_k = \begin{bmatrix} \tilde{A} \\ \tilde{B} \end{bmatrix}$, $\mathcal{T}(\tilde{A}) = \mathbb{I}_k$ e $\Phi_\mu(E) = \tilde{B}\tilde{A}^{-1}$. Lembramos também que as colunas de Q_k , assim como as de \tilde{Q}_k , formam uma base de E . Logo, existe uma matriz B de dimensão $k \times k$ tal que, $\tilde{Q}_k = Q_k B$. Desta maneira, $\Phi_{\pi\mu}(Z)$ é dado por uma sequência de aplicações diferenciáveis

$$Z \mapsto \begin{bmatrix} I_k \\ Z \end{bmatrix} = \bar{Q}R \mapsto Q_k = \Pi_\pi^{-1}\bar{Q}_k \mapsto \tilde{Q}_k = Q_k B \mapsto \Pi_\mu\tilde{Q}_k = \begin{bmatrix} \tilde{A} \\ \tilde{B} \end{bmatrix} \mapsto \tilde{B}\tilde{A}^{-1}$$

□

5

O Teorema de Johnson-Lindenstrauss

Enunciamos o teorema para familiarizar o leitor

Teorema 5.1 *Fixe $\epsilon \in (0, 1)$ e $d, n \in \mathbb{N}$. Tome $k \in \mathbb{N}$ tal que*

$$d > k \geq 4 \left(\frac{\epsilon^2}{2} - \frac{\epsilon^3}{3} \right)^{-1} \ln n.$$

Então, para qualquer conjunto V com n pontos em \mathbb{R}^d , existe um subespaço $E \subset \mathbb{R}^d$ de dimensão k para o qual temos que $\forall u, v \in V$

$$(1 - \epsilon) \|u - v\|^2 \leq \|\sqrt{d/k}(\pi_E(u) - \pi_E(v))\|^2 \leq (1 + \epsilon) \|u - v\|^2,$$

onde $\pi_E : \mathbb{R}^d \rightarrow E$ é a projeção ortogonal sobre E .

Nas primeiras duas secções vemos como o mecanismo push-forward nos permite induzir medidas uniformes em S^{d-1} e $G_k(\mathbb{R}^d)$.¹ Na seção 3, as distribuições gaussianas surgem a fim de obter um algoritmo capaz de escolher um ponto da esfera de modo uniforme. Além disso, a concentração de certas variáveis aleatórias segue estimando gaussianas. Surgem duas perspectivas: a primeira é projetar pontos aleatórios da esfera S^{d-1} em \mathbb{R}^k e a segunda, projetar um ponto fixo de S^{d-1} em subespaços aleatórios de $G_k(\mathbb{R}^d)$. Estes dois pontos de vista são equivalentes. Finalmente na seção 4 demonstramos o teorema de Johnson-Lindenstrauss. Começamos com o coração deste procedimento: uma medida invariante em $\mathcal{O}(d)$.

5.1

Medida de Haar em $\mathcal{O}(d)$

Embora a medida de Haar exista para qualquer grupo localmente compacto, restringiremos a nossa atenção a grupos compactos. A existência e unicidade da medida de Haar não serão provadas.

¹No capítulo 3, a medida uniforme em S^{d-1} é obtida normalizando a medida de Lebesgue em S^{d-1} .

Sejam G um grupo topológico compacto e μ uma medida definida nos borelianos de G . A medida μ é *invariante* se para todo boreliano $A \subset G$ e para todo $g \in G$, temos $\mu(A) = \mu\{ga : a \in A\} = \mu\{ag^{-1} : a \in A\}$.

A demonstração do seguinte teorema pode ser vista em (7, p. 78).

Teorema 5.2 (Haar) *Se G é um grupo topológico compacto, então existe uma única medida invariante μ definida nos borelianos de G com $\mu(G) = 1$. Esta medida é a medida Haar do grupo G .*

Corolário 5.3 *Seja $\mathcal{O}(d)$ o grupo ortogonal. Existe uma única medida de Haar μ em $\mathcal{O}(d)$. Ou seja, para todo boreliano $A \subset G$ e para todo $g \in \mathcal{O}(d)$, temos*

$$\mu(A) = \mu\{ga : a \in A\} = \mu\{ag^{-1} : a \in A\}. \quad (5-1)$$

Demonstração: O grupo $\mathcal{O}(d)$ é um grupo topológico compacto, assim, o resultado segue aplicando o Teorema 5.2. \square

A seguir, construiremos medidas invariantes de uma maneira geral, ou seja, em qualquer variedade. Considere $\mathcal{O}(d)$ e μ sua medida de Haar. Sejam M uma variedade e $\phi : \mathcal{O}(d) \times M \rightarrow M$ uma ação transitiva. Fixando um ponto $x \in M$ arbitrário, definimos a função $\phi_x : \mathcal{O}(d) \rightarrow M$, $g \mapsto \phi_x(g, x) = gx$. Induzimos a medida ϑ , como sendo o push-forward da medida de Haar μ . Isto é, para todo boreliano $A \subset M$

$$\vartheta(A) := \mu(\phi_x^{-1}(A)) = \mu\{g \in \mathcal{O}(d) : gx \in A\}.$$

Teorema 5.4 *A medida ϑ definida em M é a única medida de probabilidade invariante pela ação de $\mathcal{O}(d)$.²*

Demonstração: Provaremos primeiro a invariância de ϑ . Sejam $g \in \mathcal{O}(d)$ e $A \subset M$ um boreliano. Da invariância de μ , temos $\mu(\phi_x^{-1}(A)) = \mu(g\phi_x^{-1}(A))$. Se $g(\phi_x^{-1}(A)) = \phi_x^{-1}(gA)$, então $\mu(\phi_x^{-1}(A)) = \mu(g(\phi_x^{-1}(A))) = \mu(\phi_x^{-1}(gA))$. Ou seja, $\vartheta(A) = \vartheta(gA)$. Assim, a invariância de ϑ se resume a provar a igualdade $g(\phi_x^{-1}(A)) = \phi_x^{-1}(gA)$, $\forall g \in \mathcal{O}(d)$ e $\forall A \subset M$ boreliano. Veja o diagrama abaixo.

$$\begin{array}{ccc} M & \xrightarrow{g} & M \\ \phi_x^{-1} \downarrow & & \downarrow \phi_x^{-1} \\ \mathcal{O}(d) & \xrightarrow{g} & \mathcal{O}(d) \end{array} \quad \begin{array}{ccc} A & \xrightarrow{g} & gA \\ \phi_x^{-1} \downarrow & & \downarrow \phi_x^{-1} \\ \phi_x^{-1}(A) & \xrightarrow{g} & g\phi_x^{-1}(A) = \phi_x^{-1}(gA) \end{array}$$

²Lembramos ao leitor que a medida ϑ é invariante pela ação do grupo $\mathcal{O}(d)$ se para todo $g \in \mathcal{O}(d)$ e para todo conjunto boreliano $A \subset M$, temos $\vartheta(A) = \vartheta(gA)$.

Lembre que se $B \subset M$ é um conjunto qualquer, então $h \in \phi_x^{-1}(B)$ se, e somente se, $\phi(h, x) = hx \in B$. Provaremos primeiro que $g\phi_x^{-1}(A) \subset \phi_x^{-1}(gA)$. Se $\tilde{h} = gh \in g\phi_x^{-1}(A)$, com $h \in \phi_x^{-1}(A)$, i.e., $\phi(h, x) = hx \in A$, então $\phi(gh, x) = ghx \in gA$. Logo $\tilde{h} \in \phi_x^{-1}(gA)$. Reciprocamente, se $\tilde{h} \in \phi_x^{-1}(gA)$, então $\phi(\tilde{h}, x) \in gA$, i.e., existe $a \in A$ tal que $\phi(\tilde{h}, x) = \tilde{h}x = ga$. Provar que $\tilde{h} \in g\phi_x^{-1}(A)$, equivale a provar que $g^{-1}\tilde{h} \in \phi_x^{-1}(A)$, i.e., $\phi(g^{-1}\tilde{h}, x) \in A$. De fato, $g^{-1}\tilde{h}x = g^{-1}ga = a \in A$. Logo, $\phi_x^{-1}(gA) \subset g\phi_x^{-1}(A)$. Provando a invariância de ϑ .

É claro que $\mathcal{O}(d) = \phi_x^{-1}(M)$. Logo $\vartheta(M) = 1$. Assim, ϑ é uma medida de probabilidade.

Resta provar a unicidade. Suponha $\tilde{\vartheta}$ outra medida de probabilidade em M invariante pela ação do grupo $\mathcal{O}(d)$. Defina a medida $\tilde{\mu}$ em $\mathcal{O}(d)$ da seguinte maneira: se S é um boreliano de $\mathcal{O}(d)$, então

$$\tilde{\mu}(S) := \tilde{\vartheta}(\phi_x(S)) = \tilde{\vartheta}(\{\phi_x(h) : h \in S\}).$$

Lembre que a ação é transitiva, i.e., $\forall y \in M, \exists g \in \mathcal{O}(d)$ tal que $\phi_x(g) = y$. Ou seja, ϕ_x é sobrejetiva. Assim, $\phi_x(\phi_x^{-1}(A)) = A, \forall A \subset M$ boreliano e $\tilde{\mu}(\phi_x^{-1}(A)) = \tilde{\vartheta}(A)$. Mais ainda, da sobrejetividade da função ϕ_x , temos $\phi_x(\mathcal{O}(d)) = M$. Isto é $\tilde{\mu}(\mathcal{O}(d)) = 1$. Suponha $\tilde{\mu}$ invariante. Pela unicidade de μ em $\mathcal{O}(d)$, temos $\tilde{\mu} = \mu$. Portanto, $\tilde{\vartheta}(A) = \tilde{\mu}(\phi_x^{-1}(A)) = \mu(\phi_x^{-1}(A)) = \vartheta(A)$. Assim, a unicidade de $\tilde{\vartheta}$ se resume à invariância de $\tilde{\mu}$.

Novamente, da invariância de $\tilde{\vartheta}$, temos $\tilde{\vartheta}(\phi_x(S)) = \tilde{\vartheta}(g\phi_x(S))$ e supondo $g\phi_x(S) = \phi_x(gS), \forall g \in \mathcal{O}(d)$ e $\forall S \subset \mathcal{O}(d)$ subconjunto boreliano, temos $\tilde{\vartheta}(\phi_x(S)) = \tilde{\vartheta}(g\phi_x(S)) = \tilde{\vartheta}(\phi_x(gS))$. Concluindo $\tilde{\mu}(S) = \tilde{\mu}(gS)$. Logo, provar a invariância de $\tilde{\mu}$ equivale a provar $g\phi_x(S) = \phi_x(gS)$

$$\begin{array}{ccc} \mathcal{O}(d) & \xrightarrow{g} & \mathcal{O}(d) & & S & \xrightarrow{g} & S \\ \phi_x \downarrow & & \downarrow \phi_x & & \phi_x \downarrow & & \downarrow \phi_x \\ M & \xrightarrow{g} & M & & \phi_x(S) & \xrightarrow{g} & g\phi_x(S) = \phi_x(gS) \end{array}$$

Mas este fato é imediato, uma vez que $\forall h \in S \subset \mathcal{O}(d)$, temos

$$g\phi_x(h) = g\phi(h, x) = \phi(g, \phi(h, x)) = \phi(gh, x) = \phi_x(gh).$$

Portanto, a invariância de $\tilde{\mu}$ está provada. Resumindo, se existe outra medida invariante $\tilde{\vartheta}$ em M com $\tilde{\vartheta}(M) = 1$, então $\tilde{\vartheta} = \vartheta$. \square

O corolário a seguir afirma que a medida ϑ em M pode ser definida de outra maneira.

Corolário 5.5 *A medida ϑ é tal que*

$$\vartheta(A) = \mu\{g \in \mathcal{O}(d) : g^{-1}x \in A\}.$$

Demonstração: Defina $\bar{\vartheta}(A) := \mu\{g \in \mathcal{O}(d) : g^{-1}x \in A\}$. Provaremos que $\bar{\vartheta}$ é invariante pela ação do grupo ortogonal. Tome um boreliano $A \subset M$ e $h \in \mathcal{O}(d)$. Da invariância da medida de Haar μ exibida em (5-1), temos

$$\bar{\vartheta}(A) = \mu\{g \in \mathcal{O}(d) : g^{-1}x \in A\} = \mu\{gh^{-1} \in \mathcal{O}(d) : g^{-1}x \in A\}.$$

Afirmamos que $\{gh^{-1} \in \mathcal{O}(d) : g^{-1}x \in A\} = \{g' \in \mathcal{O}(d) : (g')^{-1}x \in hA\}$. De fato, se $g' = gh^{-1}$ com $g^{-1}x \in A$, então $(g')^{-1}x = (gh^{-1})^{-1}x = hg^{-1}x \in hA$. Assim, $\{gh^{-1} \in \mathcal{O}(d) : g^{-1}x \in A\} \subset \{g' \in \mathcal{O}(d) : (g')^{-1}x \in hA\}$.

Reciprocamente, tome $g' \in \mathcal{O}(d)$ tal que $(g')^{-1}x \in hA$, então $g' = (g'h)h^{-1}$ e $(g'h)^{-1}x = h^{-1}g'x \in A$. Logo, $\{gh^{-1} \in \mathcal{O}(d) : g^{-1}x \in A\} \supset \{g' \in \mathcal{O}(d) : (g')^{-1}x \in hA\}$. Portanto,

$$\bar{\vartheta}(A) = \mu\{g' \in \mathcal{O}(d) : (g')^{-1}x \in hA\} = \bar{\vartheta}(hA).$$

Isto é, $\bar{\vartheta}$ é invariante pela ação do grupo $\mathcal{O}(d)$. É claro que $\bar{\vartheta}(M) = 1$. Assim, da unicidade da medida ϑ , concluímos que $\vartheta(A) = \bar{\vartheta}(A)$. \square

5.2

Medidas invariantes em S^{d-1} e em $G_k(\mathbb{R}^d)$

Sejam $\phi : \mathcal{O}(d) \times S^{d-1} \rightarrow S^{d-1}$, $(g, x) \mapsto gx$ a ação dada no capítulo 2 e μ a medida de Haar em $\mathcal{O}(d)$. Fixemos um ponto qualquer s em S^{d-1} . Considere a aplicação $\phi_s : \mathcal{O}(d) \rightarrow S^{d-1}$, $g \mapsto gs$. Definimos a medida σ em S^{d-1} pondo para cada boreliano $A \subset S^{d-1}$

$$\sigma(A) := \mu(\phi_s^{-1}(A)) = \mu\{g \in \mathcal{O}(d) : gs \in A\}. \quad (5-2)$$

A ação ϕ é transitiva. De fato, tome a, b em S^{d-1} . A partir do vetor a , completamos uma base $\alpha = \{a_1 = a, a_2, a_3, \dots, a_d\}$ de \mathbb{R}^d e pelo processo de Gram-Schmidt, podemos garantir a ortonormalidade de α . De maneira análoga, partindo de $b \in S^{d-1}$ obtemos uma base ortonormal $\beta = \{b_1 = b, b_2, b_3, \dots, b_d\}$ de \mathbb{R}^d . Definimos o operador $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$ tal que $a_i \mapsto b_i$ para todo $i = 1, \dots, d$. Logo $g \in \mathcal{O}(d)$ e $ga = b$.

Pelo Teorema 5.4 concluímos que σ é a única medida invariante pela ação do grupo $\mathcal{O}(d)$. Além disso, pelo Corolário 5.5

$$\sigma(A) = \mu\{g \in \mathcal{O}(d) : g^{-1}s \in A\}. \quad (5-3)$$

Analogamente, induzimos uma medida em $G_k(\mathbb{R}^d)$. Sejam $\mathcal{O}(d)$ o grupo ortogonal e $G_k(\mathbb{R}^d)$ a Grassmannianas. Se $g \in \mathcal{O}(d)$ e $E \in G_k(\mathbb{R}^d)$, então gE é a imagem de E pela matriz g . Tome a ação $\phi : \mathcal{O}(d) \times G_k(\mathbb{R}^d) \rightarrow G_k(\mathbb{R}^d)$, $(g, E) \mapsto gE$ e aplicação $\phi_{\mathbb{R}^k} : \mathcal{O}(d) \rightarrow G_k(\mathbb{R}^d)$, $g \mapsto g\mathbb{R}^k$. Para todo boreliano $A \subset G_k(\mathbb{R}^d)$ definimos a medida ξ

$$\xi(A) =: \mu\{g \in \mathcal{O}(d) : g\mathbb{R}^k \in A\}. \quad (5-4)$$

A prova da transitividade de ξ é análoga ao caso da medida σ em S^{d-1} . Logo, a medida ξ é a única medida nas Grassmannianas invariante pela ação do grupo $\mathcal{O}(d)$.

5.3

Gaussianas e o lema geométrico

Dizemos que uma variável aleatória X tem *distribuição Gaussiana univariada*, $X \sim N(0, 1)$, se X tem a função densidade

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

Analogamente, um vetor aleatório $X = (X_1, X_2, \dots, X_d)$ tem *distribuição Gaussiana multivariada*, $X \sim N(0, I)$, se cada variável aleatória X_i é tal que $X_i \sim N(0, 1)$ e todas são mutuamente independentes. Neste caso, a função densidade de X é

$$f(x) = \frac{1}{(2\pi)^{d/2}} \exp\left(-\frac{\|x\|^2}{2}\right).$$

Lembramos que a distribuição de $X = (X_1, X_2, \dots, X_d)$ é uma medida de probabilidade definida nos borelianos de \mathbb{R}^d . Usando a sua função densidade, a distribuição de X é dada por

$$\Pr[X \in B] = \frac{1}{(2\pi)^{d/2}} \int_B \exp\left(-\frac{\|x\|^2}{2}\right) dx, \quad \forall B \in \mathcal{B}(\mathbb{R}^d).$$

Proposição 5.3.1 *Se $X = (X_1, X_2, \dots, X_d)$ tem distribuição Gaussiana multivariada, então sua distribuição é invariante pela ação do grupo $\mathcal{O}(d)$. Ou*

seja, $\forall B \in \mathcal{B}(\mathbb{R}^d)$ e $\forall g \in \mathcal{O}(d)$, temos

$$\Pr[X \in B] = \Pr[X \in gB]. \quad (5-5)$$

Demonstração: Faça $\alpha = \frac{1}{(2\pi)^{d/2}}$. Assim,

$$\begin{aligned} \Pr[X \in gB] &= \alpha \int_{gB} \exp\left(-\frac{\|x\|^2}{2}\right) dx = \alpha \int_B \exp\left(-\frac{\|gy\|^2}{2}\right) |\det(g'(y))| dy \\ &= \alpha \int_B \exp\left(-\frac{\|gy\|^2}{2}\right) |\det(g)| dy = \alpha \int_B \exp\left(-\frac{\|y\|^2}{2}\right) dy \\ &= \alpha \int_B \exp\left(-\frac{\|y\|^2}{2}\right) dy = \Pr[X \in B]. \quad \square \end{aligned}$$

Dizemos que um vetor aleatório Y tem *distribuição uniforme* em S^{d-1} se sua distribuição é dada por

$$\Pr[Y \in A] = \frac{\eta(A)}{\eta(S^{d-1})} = \sigma(A), \quad \forall A \in \mathcal{B}(S^{d-1}).$$

Considere a função mensurável $f : \mathbb{R}^d \setminus \{0\} \rightarrow S^{d-1}$, $x \mapsto \frac{x}{\|x\|}$ e o vetor $X = (X_1, X_2, \dots, X_d)$ com distribuição multivariada. Definimos o vetor aleatório $Y = \frac{1}{\|X\|} (X_1, X_2, \dots, X_d)$. A distribuição de Y é dada por

$$\Pr[Y \in A] = \Pr[X \in f^{-1}(A)], \quad \forall A \in \mathcal{B}(S^{d-1}). \quad (5-6)$$

O seguinte resultado afirma que Y significa escolher um ponto aleatório da esfera de modo uniforme.

Proposição 5.3.2 $Y = \frac{1}{\|X\|} (X_1, X_2, \dots, X_d)$ tem distribuição uniforme em S^{d-1} .

Demonstração: A igualdade (5-6) e a Proposição 5.3.1 implicam, respectivamente, as duas primeiras igualdades a seguir

$$\Pr[Y \in A] = \Pr[X \in f^{-1}(A)] = \Pr[X \in gf^{-1}(A)] = \Pr[X \in f^{-1}(gA)].$$

E a última igualdade é consequência de $gf^{-1}(A) = f^{-1}(gA)$. Logo, usando novamente (5-6), concluímos que $\Pr[Y \in A] = \Pr[Y \in gA]$. \square

Ao projetarmos o vetor Y em suas k primeiras coordenadas obtemos $Z = \frac{1}{\|X\|} (X_1, X_2, \dots, X_k)$. Ou seja, Z é a projeção em \mathbb{R}^k de um ponto da esfera escolhido de maneira uniforme. Seja a variável aleatória $L : S^{d-1} \rightarrow \mathbb{R}$,

$L = \|Z\|^2$. A distribuição de L na reta é dada por

$$\Pr[L \in B] = \sigma\{x \in S^{d-1} : \|\pi_{\mathbb{R}^k}(x)\|^2 \in B\}, \quad \forall B \in \mathcal{B}(\mathbb{R}), \quad (5-7)$$

Note que pela independência das coordenadas de Z e do fato $X_i \sim N(0, 1)$, temos $\mu_i = \mathbb{E} \left[\frac{X_i^2}{(X_1^2 + \dots + X_d^2)} \right] = 1/d$. Logo, $\mathbb{E}[L] = \sum_{i=1}^k \mu_i = k/d$. Mostraremos que a variável aleatória L se encontra fortemente concentrada em torno de sua média.

Lema 5.6 *Seja $k < d$. Temos então*

a. *Se $0 < \beta < 1$, então*

$$\Pr \left[L \leq \beta \frac{k}{d} \right] \leq \beta^{\frac{k}{2}} \left(1 + \frac{(1-\beta)k}{(d-k)} \right)^{\frac{(d-k)}{2}} \leq \exp \left(\frac{k}{2} (1 - \beta + \ln \beta) \right).$$

b. *Se $1 < \beta < d/k$, então*

$$\Pr \left[L \geq \beta \frac{k}{d} \right] \leq \beta^{\frac{k}{2}} \left(1 + \frac{(1-\beta)k}{(d-k)} \right)^{\frac{(d-k)}{2}} \leq \exp \left(\frac{k}{2} (1 - \beta + \ln \beta) \right).$$

Demonstração: Note que

$$L = \|Z\|^2 = \frac{X_1^2 + X_2^2 + \dots + X_k^2}{X_1^2 + X_2^2 + \dots + X_d^2}.$$

Assim,

$$\Pr \left[L \leq \beta \frac{k}{d} \right] = \Pr \left[d(X_1^2 + X_2^2 + \dots + X_k^2) \leq k\beta(X_1^2 + X_2^2 + \dots + X_d^2) \right].$$

Manipulando o lado direito desta igualdade, temos

$$\Pr \left[L \leq \beta \frac{k}{d} \right] = \Pr \left[d \sum_{i=1}^k X_i^2 \leq k\beta \sum_{i=1}^d X_i^2 \right] = \Pr \left[k\beta \sum_{i=1}^d X_i^2 - d \sum_{i=1}^k X_i^2 \geq 0 \right].$$

Para $t > 0$, temos

$$\begin{aligned}
 \Pr \left[k\beta \sum_{i=1}^d X_i^2 - d \sum_{i=1}^k X_i^2 \geq 0 \right] &= \Pr \left[\exp \left\{ t \left(k\beta \sum_{i=1}^d X_i^2 - d \sum_{i=1}^k X_i^2 \right) \right\} \geq 1 \right] \\
 &\leq \mathbb{E} \left[\exp \left\{ t \left(k\beta \sum_{i=1}^d X_i^2 - d \sum_{i=1}^k X_i^2 \right) \right\} \right] && \text{Desigualdade de Markov} \\
 &= \mathbb{E} \left[\exp \left\{ t(k\beta - d) \sum_{i=1}^k X_i^2 + tk\beta \sum_{i=k+1}^d X_i^2 \right\} \right] \\
 &= \mathbb{E} [\exp\{t(k\beta - d)X^2\}]^k \mathbb{E} [\exp\{tk\beta X^2\}]^{d-k}. && (X \sim N(0, 1))
 \end{aligned}$$

Logo

$$\Pr \left[L \leq \beta \frac{k}{d} \right] \leq \mathbb{E} [\exp\{t(k\beta - d)X^2\}]^k \mathbb{E} [\exp\{tk\beta X^2\}]^{d-k}, \quad t > 0$$

Lembremos a Proposição B.0.1 do Apêndice B: se a variável aleatória X tem distribuição normal univariada, então

$$\mathbb{E} [\exp(sX^2)] = 1/\sqrt{1-2s}, \quad -\infty < s < 1/2. \quad (5-8)$$

Logo, se $t(k\beta - d) < 1/2$ e $tk\beta < 1/2$, então

$$\mathbb{E} [\exp\{t(k\beta - d)X^2\}]^k \mathbb{E} [\exp\{tk\beta X^2\}]^{d-k} = (1-2tk\beta)^{-(d-k)/2} (1-2t(k\beta - d))^{-k/2}.$$

As restrições $t > 0$, $t(k\beta - d) < 1/2$ e $tk\beta < 1/2$ implicam $0 < t < 1/2k\beta$.

Logo, se $g(t) = (1-2tk\beta)^{-(d-k)/2} (1-2t(k\beta - d))^{-k/2}$ então

$$\Pr \left[L \leq \frac{\beta k}{d} \right] \leq g(t), \quad 0 < t < 1/2k\beta.$$

Em particular, podemos escolher $t_0 \in (0, 1/2k\beta)$ tal que $g(t_0) \leq g(t)$, $\forall t \in (0, 1/2k\beta)$. Assim, teríamos $\Pr [L \leq \frac{\beta k}{d}] \leq g(t_0)$. Nosso objetivo então é minimizar a função g . Note que

$$g(t) = \frac{1}{\sqrt{f(t)}}, \quad f(t) = (1-2tk\beta)^{(d-k)} (1-2t(k\beta - d))^k.$$

Ainda, se $f(t_0) \geq f(t)$, $\forall t \in (0, 1/2k\beta)$ então $g(t_0) \leq g(t)$, $\forall t \in (0, 1/2k\beta)$.

Logo, a fim de minimizar g , maximizamos f . Começamos derivando f e para tal chame $c_1 = (1-2tk\beta)$ e $c_2 = (1-2t(k\beta - d))$. Assim, $f'(t) = 0$ fica

$$\begin{aligned}
 f'(t) &= (d-k)c_1^{(d-k-1)}c_2^k(-2k\beta) - kc_1^{(d-k)}c_2^{(k-1)}2(d-k\beta) \\
 &= 2kc_1^{(d-k-1)}c_2^{(k-1)}\{-d\beta c_2 - \beta kc_1\} + [k\beta c_2 + dc_1] = 0
 \end{aligned}$$

Ora, observe que para $t \in (0, 1/2k\beta)$, $2kc_1^{(d-k-1)}c_2^{(k-1)} > 0$. Então

$f'(t) = 0$ implica que

$$[-d\beta c_2 - \beta k c_1] + [k\beta c_2 + d c_1] = 0.$$

Substituindo c_1 e c_2 temos

$$[-d\beta(1 - 2tk\beta + 2td) - \beta k + 2tk^2\beta^2] + [k\beta - k\beta 2t(k\beta - d) + d - 2dtk\beta] = 0.$$

Manipulando mais ainda, chegamos a $d(1 - \beta) - 2d\beta(d - k\beta)t = 0$. Assim, obtemos o único ponto crítico de f em $(0, 1/2k\beta)$

$$t_0 = \frac{1 - \beta}{2\beta(d - k\beta)}.$$

Além disso,

$$\begin{aligned} f(t_0) &= \left(1 - 2 \left(\frac{1 - \beta}{2\beta(d - k\beta)}\right) k\beta\right)^{(d-k)} \left(1 - 2 \left(\frac{1 - \beta}{2\beta(d - k\beta)}\right) (k\beta - d)\right)^k \\ &= \left(1 - \left(\frac{k - k\beta}{d - k\beta}\right)\right)^{(d-k)} \left(1 + \left(\frac{1 - \beta}{\beta}\right)\right)^k \\ &= \left(\frac{d - k}{d - k\beta}\right)^{(d-k)} \left(\frac{1}{\beta}\right)^k. \end{aligned}$$

Para provar que t_0 maximiza a função f em $(0, 1/2k\beta)$ observamos que $\lim_{t \rightarrow 0^+} f(t) = 1$ e $\lim_{t \rightarrow 1/2k\beta^-} f(t) = 0$. Estendemos f em $[0, 1/2k\beta]$ fazendo $f(0) = 1$ e $f(1/2k\beta) = 0$. Deste modo, obtemos uma função contínua definida em um compacto. Pelo Teorema de Weierstrass, a função estendida atinge seu máximo em um ponto de $[0, 1/2k\beta]$. Existe a possibilidade desse ponto ser o extremo esquerdo do intervalo $[0, 1/2k\beta]$. Descartamos esta possibilidade provando que $f(0) = 1 < f(t_0)$, onde t_0 é dado acima. Isto é, iremos provar que

$$1 < \left(\frac{d - k}{d - k\beta}\right)^{(d-k)} \left(\frac{1}{\beta}\right)^k. \quad (5-9)$$

Definimos $P(\beta) = \frac{1}{(d-k)^{d-k}} (d - \beta k)^{d-k} \beta^k$, $\forall \beta \in (0, d/k)$. Assim, a desigualdade (5-9) equivale a $P(\beta) < P(1)$. Portanto, nossa tarefa é provar que P atinge seu máximo estrito em $\beta = 1$. Derivamos P

$$P'(\beta) = \frac{1}{(d-k)^{d-k}} k\beta^{k-1} (d - \beta k)^{d-k-1} d(1 - \beta).$$

Observamos que para $0 < \beta < 1$ temos $P'(\beta) > 0$ e para $1 < \beta < d/k$, $P'(\beta) < 0$. Isto implica que se $\beta \in (0, 1)$, então P é estritamente crescente e se $\beta \in (1, d/k)$, então P é estritamente decrescente. Concluindo assim que $\beta = 1$

é um ponto de máximo estrito.

Ora, tudo isto conclui que $f(t) < f(t_0)$, $\forall t \in [0, 1/2k\beta]$. Em particular, t_0 maximiza f em $(0, 1/2k\beta)$. Finalmente, usando $g(t_0) = \frac{1}{\sqrt{f(t_0)}}$, temos que

$$\Pr \left[L \leq \beta \frac{k}{d} \right] \leq \beta^{k/2} \left(\frac{d - k\beta}{d - k} \right)^{(d-k)/2}.$$

A desigualdade $1 + x \leq \exp(x)$, $\forall x \in \mathbb{R}$ implica $\ln(1 + x) \leq x$, $\forall x > -1$. Logo, dado que $\frac{(1-\beta)k}{(d-k)} > -1$, temos,

$$\begin{aligned} \ln \left(1 + \frac{(1-\beta)k}{(d-k)} \right) &\leq \frac{(1-\beta)k}{(d-k)} \\ \frac{(d-k)}{2} \ln \left(1 + \frac{(1-\beta)k}{(d-k)} \right) &\leq \frac{k}{2}(1-\beta) \\ \frac{k}{2} \ln \beta + \ln \left(1 + \frac{(1-\beta)k}{(d-k)} \right)^{\frac{(d-k)}{2}} &\leq \frac{k}{2} + \frac{k}{2} \ln \beta - \frac{k}{2} \beta \\ \ln \beta^{\frac{k}{2}} + \ln \left(1 + \frac{(1-\beta)k}{(d-k)} \right)^{\frac{(d-k)}{2}} &\leq \frac{k}{2}(1-\beta + \ln \beta) \\ \ln \left[\beta^{\frac{k}{2}} \left(1 + \frac{(1-\beta)k}{(d-k)} \right)^{\frac{(d-k)}{2}} \right] &\leq \frac{k}{2}(1-\beta + \ln \beta) \\ \beta^{\frac{k}{2}} \left(1 + \frac{(1-\beta)k}{(d-k)} \right)^{\frac{(d-k)}{2}} &\leq \exp \left(\frac{k}{2}(1-\beta + \ln \beta) \right). \end{aligned}$$

Concluindo o primeiro item

$$\Pr \left[L \leq \beta \frac{k}{d} \right] \leq \beta^{\frac{k}{2}} \left(1 + \frac{(1-\beta)k}{(d-k)} \right)^{\frac{(d-k)}{2}} \leq \exp \left(\frac{k}{2}(1-\beta + \ln \beta) \right).$$

Agora provaremos o item *b*, onde $\beta > 1$.

$$\Pr \left[L \geq \beta \frac{k}{d} \right] = \Pr \left[d(X_1^2 + X_2^2 + \dots + X_k^2) \geq k\beta(X_1^2 + X_2^2 + \dots + X_d^2) \right].$$

Análogo ao primeiro item, manipulamos o lado direito desta equação. Assim,

$$\Pr \left[L \geq \beta \frac{k}{d} \right] \leq \mathbb{E} \left[\exp\{t(d - k\beta)X^2\} \right]^k \mathbb{E} \left[\exp\{-tk\beta X^2\} \right]^{d-k}, \quad t > 0.$$

Observe que $t > 0$, $t(d - k\beta) < 1/2$ e $-tk\beta < 1/2$ implica $0 < t < 1/2(d - k\beta)$. Logo, usando (5-8), temos

$$\mathbb{E} \left[\exp\{t(d - k\beta)X^2\} \right]^k \mathbb{E} \left[\exp\{-tk\beta X^2\} \right]^{d-k} = (1+2tk\beta)^{-(d-k)/2} (1+2t(k\beta-d))^{-k/2}$$

Portanto, $\forall t \in (0, 1/2(d - k\beta))$

$$\Pr \left[L \geq \beta \frac{k}{d} \right] \leq (1 + 2tk\beta)^{-(d-k)/2} (1 + 2t(k\beta - d))^{-k/2} = g(-t),$$

onde $g(t) = (1 - 2tk\beta)^{-(d-k)/2} (1 - 2t(k\beta - d))^{-k/2}$. Novamente, nosso propósito é minimizar $g(-t)$ em $(0, 1/2(d - k\beta))$. O procedimento é completamente análogo ao primeiro item: a fim de minimizar a função $g(-t)$ em $(0, 1/2(d - k\beta))$, fazemos

$$g(-t) = \frac{1}{\sqrt{f(t)}}, \quad f(t) = (1 + 2tk\beta)^{(d-k)} (1 + 2t(k\beta - d))^k$$

e maximizamos f em $(0, 1/2(d - k\beta))$. Começamos fazendo $f'(t) = 0$

$$\begin{aligned} f'(t) &= (d - k)c_1^{(d-k-1)}c_2^k(2k\beta) + kc_1^{(d-k)}c_2^{(k-1)}2(k\beta - d) \\ &= 2kc_1^{(d-k-1)}c_2^{(k-1)}\{[d\beta c_2 + k\beta c_1] - [k\beta c_2 + dc_1]\} = 0, \end{aligned}$$

onde $c_1 = (1 + 2tk\beta)$ e $c_2 = (1 + 2t(k\beta - d))$. Novamente, $2kc_1^{(d-k-1)}c_2^{(k-1)} \neq 0$.

Logo $[d\beta c_2 + k\beta c_1] - [k\beta c_2 + dc_1] = 0$. Obtemos assim o ponto crítico

$$t_0 = -\frac{1 - \beta}{2\beta(d - k\beta)} \in (0, 1/2(d - k\beta)).$$

O argumento para provar que t_0 maximiza a função f é análogo ao item anterior. Assim, ao substituir t_0 em $g(t_0) = \frac{1}{\sqrt{f(t_0)}}$, temos

$$\Pr \left[L \geq \beta \frac{k}{d} \right] \leq g(t_0) = \beta^{k/2} \left(\frac{d - k\beta}{d - k} \right)^{(d-k)/2}.$$

A restrição $\beta < d/k$ implica que $\frac{(1-\beta)k}{(d-k)} > -1$. Logo, a desigualdade $\beta^{k/2} \left(\frac{d-k\beta}{d-k} \right)^{(d-k)/2} \leq \exp\left(\frac{k}{2}(1 - \beta + \ln \beta)\right)$ também é igual ao item anterior.

Logo,

$$\Pr \left[L \geq \beta \frac{k}{d} \right] \leq \beta^{\frac{k}{2}} \left(1 + \frac{(1 - \beta)k}{(d - k)} \right)^{\frac{(d-k)}{2}} \leq \exp \left(\frac{k}{2}(1 - \beta + \ln \beta) \right).$$

Concluindo o segundo item. □

Fixe $s \in S^{d-1}$ qualquer. Defina a variável aleatória $\mathcal{L} : G_k(\mathbb{R}^d) \rightarrow \mathbb{R}$, $E \mapsto \|\pi_E(s)\|^2$. A distribuição de \mathcal{L} na reta é dada por

$$\Pr[\mathcal{L} \in B] = \xi\{E \in G_k(\mathbb{R}^d) : \|\pi_E(s)\|^2 \in B\}, \quad \forall B \in \mathcal{B}(\mathbb{R}). \quad (5-10)$$

Fato: $\forall s \in S^{d-1}$ e $\forall g \in \mathcal{O}(d)$, temos

$$\|\pi_{g\mathbb{R}^k}(s)\| = \|\pi_{\mathbb{R}^k}(g^{-1}s)\|. \quad (5-11)$$

Demonstração: Sejam $\{e_1, \dots, e_k\}$ a base canônica de \mathbb{R}^k e $\{ge_1, \dots, ge_k\}$ uma base ortonormal de $g\mathbb{R}^k$. Assim, observamos que $\pi_{g\mathbb{R}^k}(s) = \sum_{i=1}^k \langle s, ge_i \rangle ge_i$ e $\pi_{\mathbb{R}^k}(g^{-1}s) = \sum_{i=1}^k \langle (g^{-1}s), e_i \rangle e_i$. Logo, pelo teorema de Pitágoras, temos

$$\|\pi_{g\mathbb{R}^k}(s)\|^2 = \sum_{i=1}^k \langle s, ge_i \rangle^2 \quad \text{e} \quad \|\pi_{\mathbb{R}^k}(g^{-1}s)\|^2 = \sum_{i=1}^k \langle g^{-1}s, e_i \rangle^2.$$

Além disso, $\langle s, ge_i \rangle = \langle g^*s, e_i \rangle = \langle g^{-1}s, e_i \rangle$. Portanto, concluímos que $\|\pi_{g\mathbb{R}^k}(s)\|^2 = \|\pi_{\mathbb{R}^k}(g^{-1}s)\|^2$. \square

Teorema 5.7 *A variável aleatória $L = \|Z\|^2$, definida no Lema 5.6, possui a mesma distribuição de \mathcal{L} .*

Demonstração: Seja $B \in \mathcal{B}(\mathbb{R})$ um boreliano da reta. Por definição temos que $\Pr[L \in B] = \sigma\{x \in S^{d-1} : \|\pi_{\mathbb{R}^k}(x)\|^2 \in B\}$ ³. Por outro lado, σ é o push-forward da medida de Haar do grupo $\mathcal{O}(d)$ ⁴, portanto

$$\sigma\{x \in S^{d-1} : \|\pi_{\mathbb{R}^k}(x)\|^2 \in B\} = \mu\{g \in \mathcal{O}(d) : \|gs_{\mathbb{R}^k}\|^2 \in B\}.$$

Além disso, a igualdade 5-3, implica que

$$\sigma\{x \in S^{d-1} : \|\pi_{\mathbb{R}^k}(x)\|^2 \in B\} = \mu\{g \in \mathcal{O}(d) : \|\pi_{\mathbb{R}^k}(g^{-1}s)\|^2 \in B\}.$$

Assim, concluímos que

$$\Pr[L \in B] = \mu\{g \in \mathcal{O}(d) : \|\pi_{\mathbb{R}^k}(g^{-1}s)\|^2 \in B\}. \quad (5-12)$$

Logo,

$$\begin{aligned} \Pr[\mathcal{L} \in B] &\stackrel{(5-10)}{=} \xi\{E \in G_k(\mathbb{R}^d) : \|\pi_E(s)\|^2 \in B\} \\ &\stackrel{(5-4)}{=} \mu\{g \in \mathcal{O}(d) : \|\pi_{g\mathbb{R}^k}(s)\|^2 \in B\} \\ &\stackrel{(5-11)}{=} \mu\{g \in \mathcal{O}(d) : \|\pi_{\mathbb{R}^k}(g^{-1}s)\|^2 \in B\} \stackrel{(5-12)}{=} \Pr[L \in B]. \end{aligned}$$

\square

O corolário a seguir é uma consequência direta do Teorema 5.7 e do Lema 5.6.

³Veja (5-7)

⁴Em (5-2) definimos $\sigma(A) := \mu(\phi_s^{-1}(A)) = \mu\{g \in \mathcal{O}(d) : gs \in A\}$

Corolário 5.8 *Seja $k < d$. Temos então:*

a. *Se $0 < \beta < 1$, então*

$$\Pr \left[\mathcal{L} \leq \frac{\beta k}{d} \right] \leq \exp \left(\frac{k}{2} (1 - \beta + \ln \beta) \right).$$

b. *Se $1 < \beta < d/k$, então*

$$\Pr \left[\mathcal{L} \geq \frac{\beta k}{d} \right] \leq \exp \left(\frac{k}{2} (1 - \beta + \ln \beta) \right).$$

5.4

Demonstração do teorema de Johnson-Lindenstrauss

Teorema. (Johnson-Lindenstrauss) Fixe $\epsilon \in (0, 1)$ e $d, n \in \mathbb{N}$. Tome $k \in \mathbb{N}$ tal que

$$d > k \geq 4 \left(\frac{\epsilon^2}{2} - \frac{\epsilon^3}{3} \right)^{-1} \ln n.$$

Então, para qualquer conjunto V com n pontos em \mathbb{R}^d , existe um subespaço $E \subset \mathbb{R}^d$ de dimensão k para o qual temos que $\forall u, v \in V$

$$(1 - \epsilon) \|u - v\|^2 \leq \|\sqrt{d/k}(\pi_E(u) - \pi_E(v))\|^2 \leq (1 + \epsilon) \|u - v\|^2,$$

onde $\pi_E : \mathbb{R}^d \rightarrow E$ é a projeção ortogonal sobre E .

Demonstração: Para cada $E \in G_k(\mathbb{R}^d)$ considere a projeção $\pi_E : \mathbb{R}^d \rightarrow E$, $x \mapsto \pi_E(x)$. Fixamos $u, v \in V$ e fazemos $s = \frac{u-v}{\|u-v\|} \in S^{d-1}$. Assim, definimos a variável aleatória $\mathcal{L} : G_k(\mathbb{R}^d) \rightarrow \mathbb{R}$, $E \mapsto \|\pi_E(s)\|^2$, $\pi_E(s) = \frac{\pi_E(u) - \pi_E(v)}{\|u-v\|}$.

Tomamos $\beta_1 = 1 - \epsilon$ e aplicamos o item a do Corolário 5.8

$$\begin{aligned} \xi \left[\frac{\|\sqrt{d/k}(\pi_E(u) - \pi_E(v))\|^2}{\|u-v\|^2} \leq \beta \right] &= \Pr \left[\mathcal{L} \leq \frac{\beta k}{d} \right] \leq \exp \left(\frac{k}{2} (1 - \beta + \ln \beta) \right) \\ &= \exp \left(\frac{k}{2} (1 - (1 - \epsilon) + \ln(1 - \epsilon)) \right) \\ &= \exp \left(\frac{k}{2} (\epsilon + \ln(1 - \epsilon)) \right). \end{aligned}$$

De $\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n$, concluímos que $\ln(1-x) \leq -x - \frac{x^2}{2}$ para todo $0 \leq x < 1$. Logo, $\exp \left(\frac{k}{2} (\epsilon + \ln(1 - \epsilon)) \right) \leq \exp \left(\frac{k}{2} (\epsilon - \epsilon - \frac{\epsilon^2}{2}) \right) = \exp \left(\frac{-k\epsilon^2}{4} \right)$. Manipulando a hipótese $k \geq 4 \left(\frac{\epsilon^2}{2} - \frac{\epsilon^3}{3} \right)^{-1} \ln n$, obtemos $\frac{-k\epsilon^2}{4} \leq -2 \ln n$. Logo,

$$\xi \left[\frac{\|\sqrt{d/k}(\pi_E(u) - \pi_E(v))\|^2}{\|u - v\|^2} \leq \beta \right] \leq \exp(\ln n^{-2}) = \frac{1}{n^2}.$$

Analogamente, se tomamos $\beta_2 = 1 + \epsilon$ e aplicamos o item *b* do Corolário 5.8, temos

$$\begin{aligned} \xi \left[\frac{\|\sqrt{d/k}(\pi_E(u) - \pi_E(v))\|^2}{\|u - v\|^2} \geq \beta \right] &= \Pr \left[\mathcal{L} \geq \frac{\beta k}{d} \right] \leq \exp \left(\frac{k}{2} (1 - \beta_2 + \ln \beta_2) \right) \\ &= \exp \left(\frac{k}{2} (1 - (1 + \epsilon) + \ln(1 + \epsilon)) \right) \\ &= \exp \left(\frac{k}{2} (-\epsilon + \ln(1 + \epsilon)) \right). \end{aligned}$$

Por Taylor, $\ln(1 + x) \leq x - \frac{x^2}{2} + \frac{x^3}{3}$ para todo $x \geq 0$. Logo, $\exp \left(\frac{k}{2} (-\epsilon + \ln(1 + \epsilon)) \right) \leq \exp \left(\frac{k}{2} (-\epsilon + \epsilon - \frac{\epsilon^2}{2} + \frac{\epsilon^3}{3}) \right) = \exp \left(-\frac{k}{2} \left(\frac{\epsilon^2}{2} - \frac{\epsilon^3}{3} \right) \right)$. Novamente, manipulando $k \geq 4 \left(\frac{\epsilon^2}{2} - \frac{\epsilon^3}{3} \right)^{-1} \ln n$, temos $-\frac{k}{2} \left(\frac{\epsilon^2}{2} - \frac{\epsilon^3}{3} \right) \leq -2 \ln n$. Logo,

$$\xi \left[\frac{\|\sqrt{d/k}(\pi_E(u) - \pi_E(v))\|^2}{\|u - v\|^2} \geq \beta \right] \leq \exp(\ln n^{-2}) = \frac{1}{n^2}.$$

Assim,

$$\xi \left\{ E \in G_k(\mathbb{R}^d) : \frac{\|\sqrt{d/k}(\pi_E(u) - \pi_E(v))\|^2}{\|u - v\|^2} \notin (1 - \epsilon, 1 + \epsilon) \right\} \leq \frac{2}{n^2}.$$

Portanto, a probabilidade de $\frac{\|\sqrt{d/k}(\pi_E(u) - \pi_E(v))\|^2}{\|u - v\|^2}$ não pertencer ao intervalo $(1 - \epsilon, 1 + \epsilon)$ para dois pontos quaisquer $u, v \in V$ é menor ou igual a

$$\binom{n}{2} \times \frac{2}{n^2} = \frac{n!}{(n-2)!2} \times \frac{2}{n^2} = \frac{n(n-1)}{n^2} = 1 - \frac{1}{n}.$$

Logo, concluímos que para qualquer para de vetores $u, v \in V$, temos

$$\begin{aligned} &\xi \left\{ E \in G_k(\mathbb{R}^d) : (1 - \epsilon) \leq \frac{\|\sqrt{d/k}(\pi_E(u) - \pi_E(v))\|^2}{\|u - v\|^2} \leq (1 + \epsilon) \right\} \\ &= 1 - \xi \left\{ E \in G_k(\mathbb{R}^d) : \frac{\|\sqrt{d/k}(\pi_E(u) - \pi_E(v))\|^2}{\|u - v\|^2} \notin (1 - \epsilon, 1 + \epsilon) \right\} \\ &\geq \frac{1}{n}. \end{aligned}$$

Ou seja, dadas todas as condições suficientes do teorema, existe $\pi_E : \mathbb{R}^d \rightarrow E$ tal que

$$(1 - \epsilon)\|u - v\|^2 \leq \|\sqrt{d/k}(\pi_E(u) - \pi_E(v))\|^2 \leq (1 + \epsilon)\|u - v\|^2.$$

□

A

Noções de Variedades

Lembramos para conveniência do leitor o vocabulário básico associado à construção de variedades. Seja M um conjunto qualquer. Uma *carta local* (U, φ) é uma bijeção $\varphi : U \rightarrow \varphi(U)$ de um subconjunto $U \subset M$ sobre um aberto $\varphi(U) \subset \mathbb{R}^m$. Um *atlas* \mathfrak{A} de dimensão m e diferenciável de classe C^k ($k > 1$) sobre M é uma família de cartas locais $\{(U_i, \varphi_i) : i \in I, \varphi_i(U_i) \subset \mathbb{R}^m\}$ que cobrem M , no sentido que $M = \cup_{i \in I} U_i$ e para as quais todas as mudanças de coordenadas são de classe C^k . Ou seja, se (U_i, φ_i) e (U_j, φ_j) são duas cartas de \mathfrak{A} com $U_i \cap U_j \neq \emptyset$, então a aplicação

$$\varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \rightarrow \varphi_j(U_i \cap U_j)$$

é um difeomorfismo de classe C^k .

Uma carta local (U, φ) é *admissível* a um atlas \mathfrak{A} se $\mathfrak{A} \cup (U, \varphi)$ ainda é um atlas de classe C^k em M .

Uma *variedade diferenciável* de dimensão m e classe C^k é um par ordenado (M, \mathfrak{A}) onde M é um conjunto e \mathfrak{A} é um atlas de dimensão m e de classe C^k sobre M . Por praticidade escrevemos variedade diferenciável M ou simplesmente variedade M quando o atlas for fácil de identificar. Quando a estrutura é C^∞ , diremos que a variedade é suave.

Seja M uma variedade diferenciável M . Um subconjunto $A \subset M$ é aberto se para cada $a \in A$ existe uma carta local admissível (U, φ) tal que $a \in U$ e $U \subset A$. Desta maneira, induzimos uma topologia à variedade M .

Seja (M, \mathfrak{A}) uma variedade de dimensão m . Um subconjunto $N \subset M$ é uma *subvariedade* de M e dimensão $n \leq m$ se para todo $x \in N$ existe uma carta local (U, φ) admissível ao atlas \mathfrak{A} com $x \in U$ e tal que $\varphi : U \rightarrow \mathbb{R}^n \times \mathbb{R}^{m-n}$ satisfaz $\varphi(U \cap N) = \varphi(U) \cap (\mathbb{R}^n \times \{0\})$.

Seja $U_0 \subset \mathbb{R}^m$ um aberto. Uma aplicação diferenciável $f : U_0 \rightarrow \mathbb{R}^{m+n}$ é uma *imersão* quando, para cada $x \in U_0$, $f'(x) : \mathbb{R}^m \rightarrow \mathbb{R}^{m+n}$ é injetiva. Uma *parametrização* de classe C^k e dimensão m de um conjunto $U \subset \mathbb{R}^n$ é uma imersão $\varphi : U_0 \rightarrow U$ de classe C^k que é, ao mesmo tempo, um homeomorfismo do aberto $U_0 \subset \mathbb{R}^m$ sobre U .

Um conjunto $M \subset \mathbb{R}^n$ chama-se uma *superfície* de dimensão m e classe C^k quando todo ponto $p \in M$ está contido em algum aberto $U \subset \mathbb{R}^n$ de maneira que $V = U \cap M$ é imagem de uma parametrização $\varphi : V_0 \subset \mathbb{R}^m \rightarrow V$, de dimensão m e classe C^k .

B

Noções de Probabilidade

Um *espaço de probabilidade* é um espaço de medida $(\Omega, \mathcal{A}, \xi)$ onde \mathcal{A} é uma σ -álgebra de eventos aleatórios e ξ é uma medida de probabilidade, ou seja, $\xi(X) = 1$. Uma *variável aleatória* em um espaço de probabilidade $(\Omega, \mathcal{A}, \xi)$ é uma função real $X : \Omega \rightarrow \mathbb{R}$ tal que $[X \in B] \in \mathcal{A}, \forall B \in \mathcal{B}(\mathbb{R})$. Ou seja, $[X \in B]$ é um evento aleatório para todo boreliano $B \subset \mathbb{R}$. Se $f : \mathbb{R} \rightarrow \mathbb{R}$ é uma função mensurável, então $X \circ f$ é ainda uma variável aleatória.

Se X é uma variável aleatória em um espaço de medida $(\Omega, \mathcal{A}, \xi)$, então a medida de probabilidade nos borelianos de \mathbb{R} dada por

$$\Pr(B) = \xi\{\omega \in \Omega : X(\omega) \in B\}, \quad B \in \mathcal{B}(\mathbb{R})$$

é chamada de *distribuição de probabilidade* ou simplesmente *distribuição* de X em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Um vetor $X = (X_1, X_2, \dots, X_d)$ onde cada coordenada é uma variável aleatória e todas elas definidas no mesmo espaço de probabilidade $(\Omega, \mathcal{A}, \xi)$ é chamado de *vetor aleatório*. Analogamente, definimos a distribuição de um vetor aleatório $X = (X_1, X_2, \dots, X_d)$ como

$$\Pr(B) = \xi\{\omega \in \Omega : X(\omega) \in B\}, \quad B \in \mathcal{B}(\mathbb{R}^d).$$

Seja X uma variável aleatória definida no espaço de medida $(\Omega, \mathcal{A}, \xi)$. Na linguagem de teoria da medida, X é uma função mensurável e a *esperança* X é definida como sendo a integral com respeito à medida ξ . A esperança de X é denotada como

$$\mathbb{E}[X] := \int_{\Omega} X d\xi.$$

Se X é uma variável aleatória não negativa e $a > 0$, então a *desigualdade de Markov* afirma que

$$\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.$$

Terminamos este apêndice como a seguinte proposição:

Proposição B.0.1 *Se a variável aleatória X tem distribuição normal univariada, então para $-\infty < s < 1/2$, temos $\mathbb{E}[\exp(sX^2)] = 1/\sqrt{1-2s}$.*

Demonstração: Chame de $\alpha = \frac{1}{\sqrt{2\pi}}$

$$\begin{aligned}\mathbb{E} [\exp (sX^2)] &= \alpha \int_{\mathbb{R}} \exp (st^2) \exp (-t^2/2) dt = \alpha \int_{\mathbb{R}} \exp ((2st^2 - t^2)/2) dt \\ &= \alpha \int_{\mathbb{R}} \exp ((t(\sqrt{1-2s}))^2/2) dt.\end{aligned}$$

Fazendo $z = t(\sqrt{1-2s})$ e assumindo $-\infty < s < 1/2$, concluímos que

$$\mathbb{E} [\exp (sX^2)] = \frac{1}{\sqrt{1-2s}} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp (z^2/2) dz = 1/\sqrt{1-2s}. \quad (\text{B-1})$$

□

Referências Bibliográficas

- [1] ABRAHAM, R.; MARSDEN, J. E.; RATIU, T. S. ; CUSHMAN, R.. **Foundations of mechanics**. Benjamin/Cummings Publishing Company, 1978.
- [2] BARTLE, R. G.; BARTLE, R. G.. **The elements of integration and Lebesgue measure**. Wiley Online Library, 1995.
- [3] DASGUPTA, S.; GUPTA, A.. **An elementary proof of a theorem of johnson and lindenstrauss**. Random Structures & Algorithms, 22(1):60–65, 2002.
- [4] FOLLAND, G. B.; FOLLAND, G.. **Real analysis: modern techniques and their applications**, volumen 2. Wiley New York, 1984.
- [5] HERSTEIN, I. N.; HERSTEIN, I.. **Abstract algebra**, volumen 21990. Macmillan New York, 1986.
- [6] JAMES, B. R.. **Probabilidade: Um curso em nível intermediário, coleção euclides**. Rio de Janeiro. IMPA, 2ª Edição, 1996.
- [7] KRANTZ, S. G.; PARKS, H. R.. **Geometric integration theory**. Birkhäuser Boston, 2008.
- [8] LEDOUX, M.. **The concentration of measure phenomenon**, volumen 89. Amer Mathematical Society, 2001.
- [9] LIMA, E. L.. **Algebra linear, sétima edição**. SBM, Coleção Matemática Universitária, 2004.
- [10] LIMA, E. L.. **Curso de análise vol. 2 (oitava edição)**. Projeto Euclides. IMPA, 2005.
- [11] LIMA, E. L.. **Variedades diferenciáveis**. Publicações Matemáticas, 2007.
- [12] MUNKRES, J. R.. **Topology**. Prentice Hall, 2nd edition, 2000.
- [13] ROYDEN, H.. **Real analysis**. Macmillan New York, 1968.

- [14] SHIRYAEV, A.. **Probability. number 95 in graduate texts in mathematics**, 1996.
- [15] TREFETHEN, L. N.; BAU III, D.. **Numerical linear algebra**. Número 50. Society for Industrial Mathematics, 1997.