Notas Para um Curso de Cálculo Avançado

Daniel V. Tausk

Sumário

CAPÍTULO 1

Diferenciação

1.1. Notação em Cálculo Diferencial

Nesta seção preliminar quero esclarecer alguns aspectos da notação que é normalmente usada em cursos elementares de Cálculo Diferencial. Na discussão que segue, farei menções a conceitos (como diferenciação e integração) que serão definidos ao longo do restante destas notas; em particular, estarei falando sobre certos conceitos *antes* de tê-los definido. Isso não é tão estranho quanto parece: em primeiro lugar, como essa discussão preliminar (um pouco informal) gira apenas em torno de *notação* e não de *conteúdo*, não faz realmente muita diferença o significado dos conceitos. Em segundo lugar, é bem provável que o leitor típico destas notas já tenha feito cursos básicos de Cálculo Diferencial (embora isso não seja estritamente necessário).

A motivação para escrever estes esclarecimentos iniciais é proveniente do fato que é comum encontrar alguns abusos de notação em cursos de Cálculo (principalmente quando são direcionados a não-matemáticos) e também pela minha sensação de que muitas coisas sobre essa notação não costumam ser devidamente explicadas. Só para dar um exemplo, observe que as expressões:

$$\frac{\partial f}{\partial x}(x^2y, y-x), \tag{1.1.1}$$

$$\frac{\partial}{\partial x}\, f(x^2y, y-x) \tag{1.1.2}$$

têm significados bastante diferentes, embora tenham aparência similar. As ocorrências de "$\frac{\partial}{\partial x}$" em ambas as expressões têm funções diferentes. Como eu nunca vi alguém fazer uma discussão sobre o assunto, achei que seria adequado incluí-la neste curso.

1.1.1. Uma distinção fundamental: nomes para funções *versus* fórmulas para funções. Parte da confusão que ocorre em cursos de Cálculo é devida à falta da consciência da diferença entre um *nome* para uma função e uma *fórmula* para uma função. Vamos elucidar a questão. Quando dizemos:

"Seja $f:\mathbb{R}\to\mathbb{R}$ a função definida por $f(x)=x^3-x\cos(x^2)$, para todo $x\in\mathbb{R}$."

nós estamos descrevendo uma função e avisando que ela será denotada pela letra f. A letra f é um *nome* para a função que foi descrita. Quando trabalhamos com Cálculo nós normalmente precisamos falar sobre uma quantidade enorme de funções que aparecem ao longo de longas manipulações algébricas; por exemplo, numa manipulação rotineira como:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left[\cos(x^3+x^2) - x^2\operatorname{sen}(x^5)\right] &= -\operatorname{sen}(x^3+x^2)\,\frac{\mathrm{d}}{\mathrm{d}x}\,(x^3+x^2)\\ &\quad - \Big(\frac{\mathrm{d}}{\mathrm{d}x}\,x^2\Big)\operatorname{sen}x - x^2\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{sen}(x^5)\\ &= -\operatorname{sen}(x^3+x^2)(3x^2+2x)\\ &\quad - 2x\operatorname{sen}x - 5x^6\cos(x^5),\end{aligned}$$

nós precisamos fazer referência às derivadas das funções f, g, h, k definidas por:

$$f(x) = \cos(x^3+x^2) - x^2\operatorname{sen}(x^5), \quad g(x) = x^3+x^2,$$
$$h(x) = x^2, \quad k(x) = \operatorname{sen}(x^5),$$

para todo $x \in \mathbb{R}$. Se só houvesse notação para falar sobre a derivada de uma função através do *nome* dessa função os cálculos acima ficariam assim:

$$\begin{aligned}f'(x) &= -\operatorname{sen}(x^3+x^2)g'(x) - h'(x)\operatorname{sen}x - x^2k'(x)\\ &= -\operatorname{sen}(x^3+x^2)(3x^2+2x) - 2x\operatorname{sen}x - 5x^6\cos(x^5).\end{aligned}$$

Como é muito desagradável e desajeitado dar nomes para todas as funções que aparecem ao longo das manipulações algébricas que fazemos quando trabalhamos com Cálculo, nós usamos uma estratégia prática para fazer referência a uma função sem que ela tenha um nome: nós usamos frases como "considere a função $x^3+\operatorname{sen}(x^2)$", "tome a derivada da função $\cos(x+x^3)$", etc; mas $x^3+\operatorname{sen}(x^2)$ e $\cos(x+x^3)$ não são *nomes* para funções e sim os resultados obtidos pela avaliação de funções num ponto x de seus domínios. Expressões tais como $x^3+\operatorname{sen}(x^2)$ e $\cos(x+x^3)$ serão chamadas *fórmulas* para funções. Antes de continuarmos, deve-se esclarecer que fórmulas não identificam funções de forma inequívoca como os nomes fazem. Para começar, uma fórmula não deixa claro qual seja o domínio e o contra-domínio da função que se quer considerar[1], mas isso em geral é um problema menor. Um problema mais sério é o seguinte; considere a fórmula:

$$\operatorname{sen}(xy) - y\cos(x+1). \tag{1.1.3}$$

[1]Embora exista normalmente um "domínio natural" para uma fórmula, isto é, o maior conjunto no qual faz sentido calcular a fórmula, às vezes nós podemos preferir considerar uma função cujo domínio é menor do que o "domínio natural" de uma fórmula para essa função. Por exemplo, nós poderíamos considerar a função $f :]0,+\infty[\to \mathbb{R}$ definida por $f(x) = \frac{1}{x}$, para todo $x > 0$, embora a fórmula $\frac{1}{x}$ faça sentido para todo $x \neq 0$.

A que função ela se refere? Pode ser que $y \in \mathbb{R}$ seja fixado e estejamos interessados em fazer referência à função[2]:

$$\mathbb{R} \ni x \longmapsto \operatorname{sen}(xy) - y\cos(x+1) \in \mathbb{R},$$

mas pode ser que $x \in \mathbb{R}$ seja fixado e estamos fazendo referência à função:

$$\mathbb{R} \ni y \longmapsto \operatorname{sen}(xy) - y\cos(x+1) \in \mathbb{R}.$$

Pode ser também que estejamos pensando numa função definida em $\mathbb{R}^2$:

$$\mathbb{R}^2 \ni (x,y) \longmapsto \operatorname{sen}(xy) - y\cos(x+1) \in \mathbb{R},$$

ou, como não há regra alguma que diga que x é a "primeira variável" e y a "segunda variável" (imagine se usássemos as letras λ e t no lugar de x e y! qual vêm primeiro?) nós poderíamos também estar fazendo referência à função:

$$\mathbb{R}^2 \ni (y,x) \longmapsto \operatorname{sen}(xy) - y\cos(x+1) \in \mathbb{R}.$$

É necessário então (principalmente quando se lida com fórmulas que contém várias variáveis[3]) que seja fornecida alguma informação adicional para que um leitor possa identificar a que função uma dada fórmula se refere.

Uma estratégia para se tornar mais claro e rigoroso um texto em que fórmulas de funções são usadas onde se deveriam usar nomes de funções seria a de se introduzir uma simbologia que nos permitisse transformar facilmente fórmulas de funções em nomes de funções. Por exemplo, usar (como já fizemos acima) a expressão:

$$A \ni x \mapsto \langle \text{fórmula} \rangle \in B$$

ou, mais abreviadamente, $x \mapsto \langle \text{fórmula} \rangle$ como um nome para a função $f : A \to B$ tal que $f(x) = \langle \text{fórmula} \rangle$, para todo $x \in A$. Assim, por exemplo, se $\mathbb{R} \ni x \mapsto x^3 \in \mathbb{R}$ (ou, mais abreviadamente, $x \mapsto x^3$) seria um nome para a função $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = x^3$, para todo $x \in \mathbb{R}$; poderíamos escrever[4] então $(x \mapsto x^3)(2) = 8$, já que $(x \mapsto x^3)(2)$ é o mesmo que $f(2)$. Poderíamos escrever também $(x \mapsto x^3)'(2) = 12$, já que $(x \mapsto x^3)'(2)$ é o mesmo que $f'(2)$. Evitarei essa abordagem, pois não quero introduzir notações que sejam muito diferentes das que aparecem normalmente nos livros de Cálculo. A estratégia que adotarei — que é a mesma que os livros de Cálculo adotam, só que sem uma explicação explícita — será a de usar símbolos para derivadas (e também limites ou integrais) que possam ser usados junto com *fórmulas* de funções e não só com *nomes* de funções,

[2]No que segue usarei a expressão "$A \ni x \mapsto f(x) \in B$" como *nome* para uma função $f : A \to B$. Assim, por exemplo, a função $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = x^2$, para todo $x \in \mathbb{R}$, será denotada por $\mathbb{R} \ni x \mapsto x^2 \in \mathbb{R}$. Repare na diferença entre os símbolos "$\mapsto$" e "$\longrightarrow$".

[3]Mas não somente nesse caso. Por exemplo, a fórmula x^2 poderia fazer referência à função $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x,y) = x^2$, para todos $x, y \in \mathbb{R}$.

[4]Daniel J. Bernstein faz exatamente isso em [**1**].

como é o caso da notação f' para a derivada da função cujo *nome* é f. A notação para derivadas usando fórmulas de funções é a seguinte:

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\,\langle\text{fórmula}\rangle, \tag{1.1.4}$$

onde o símbolo deve ser trocado por uma variável e ⟨fórmula⟩ por uma fórmula que tipicamente (mas não necessariamente) utiliza a variável que será substituída por α. A expressão (1.1.4) significa $f'(\alpha)$, onde f é a função $\alpha \mapsto \langle\text{fórmula}\rangle$. Assim, temos por exemplo:

$$\frac{\mathrm{d}}{\mathrm{d}x}\,(x^3+x^2) = 3x^2+2x, \quad \frac{\mathrm{d}}{\mathrm{d}t}\cos(t^5) = -5t^4\,\mathrm{sen}(t^5),$$
$$\frac{\mathrm{d}}{\mathrm{d}y}\,(x^2+xy+y^2) = x+2y.$$

Note que *não devemos* escrever $(x^3+x^2)'$ pois o apóstrofe (ou "linha") deve ser usado apenas junto ao *nome* de uma função e não junto a uma fórmula. Note que uma expressão como $(x^2+xy+y^2)'$ seria realmente ambígua, já que não fica claro qual é a função que se pretende derivar. Por outro lado, não devemos escrever $\frac{\mathrm{d}}{\mathrm{d}x}f$ ou $\frac{\mathrm{d}f}{\mathrm{d}x}$, já que $\frac{\mathrm{d}}{\mathrm{d}x}$ deve ser usado apenas junto à *fórmula* de uma função e não junto ao nome (veja, no entanto, Subseção 1.1.2 a seguir para uma explicação sobre $\frac{\mathrm{d}f}{\mathrm{d}x}$). No entanto, a expressão:

$$\frac{\mathrm{d}}{\mathrm{d}x}f(x)$$

é perfeitamente correta e é igual a $f'(x)$, já que $f(x)$ é uma fórmula para a função f.

Alguns autores podem preferir utilizar ∂ no lugar de d em (1.1.4), de modo que:

$$\frac{\partial}{\partial\alpha}\,\langle\text{fórmula}\rangle$$

tem exatamente o mesmo significado que (1.1.4). Eu, no entanto, utilizarei sempre a forma (1.1.4). O uso de ∂ para derivadas parciais ou direcionais é discutido na Subseção 1.1.2 a seguir.

Note que uma expressão como (1.1.4) não é um *nome* para uma função, mas sim uma *fórmula* (que poderia corresponder a várias funções, especialmente no caso em que temos várias variáveis). *Não podemos* então escrever:

$$\Big(\frac{\mathrm{d}}{\mathrm{d}x}\,(x^3+x^2)\Big)(1) = 5;$$

de fato, para avaliar uma função num ponto x_0 deve-se colocar o símbolo (x_0) na frente de um *nome* para a função em questão (veja, por exemplo, que não escrevemos $(x^2+x^3)(1)$, para avaliar a função $x \mapsto x^2+x^3$ no ponto 1). Para que possamos avaliar uma função num ponto sem que tenhamos um nome para a mesma, é conveniente utilizar a notação:

$$\langle\text{fórmula}\rangle|_{\alpha=\alpha_0} \tag{1.1.5}$$

onde α deve ser substituída por uma variável e α_0 por uma fórmula. A expressão (1.1.5) significa $f(\alpha_0)$, onde f é a função $\alpha \mapsto \langle\text{fórmula}\rangle$. Por exemplo, temos:

$$(x^2 + x^3)|_{x=1} = 1^2 + 1^3 = 2, \quad (x^2 + xy + y^2)|_{y=3} = x^2 + 3x + 9,$$

ou ainda (estendendo um pouco a notação que acabamos de introduzir):

$$(x^2 + xy + y^2)|_{(x,y)=(2,3)} = 2^2 + 2 \cdot 3 + 3^2 = 19.$$

Dessa forma, temos:

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} \langle\text{fórmula}\rangle\Big|_{\alpha=\alpha_0} = f'(\alpha_0),$$

onde f é a função $\alpha \mapsto \langle\text{fórmula}\rangle$; por exemplo, temos:

$$\frac{\mathrm{d}}{\mathrm{d}x} (x^3 + x^2)\Big|_{x=1} = (3x^2 + 2x)|_{x=1} = 5.$$

A distinção entre nomes e fórmulas nos permite também entender melhor a notação para integrais. Por exemplo, se f é uma função a valores reais cujo domínio contém o intervalo[5] $[a, b]$, podemos denotar a integral de f no intervalo $[a, b]$ (ou seja, a integral da restrição de f a $[a, b]$) por:

$$\int_a^b f.$$

No entanto, nós *não escrevemos* $\int_a^b x^2 = \frac{1}{3}(b^3 - a^3)$, já que x^2 não é o nome de uma função. Se queremos escrever a integral de uma função para a qual não demos nome, devemos usar a notação:

$$\int_a^b \langle\text{fórmula}\rangle \,\mathrm{d}\alpha, \tag{1.1.6}$$

onde α deve ser substituído por uma variável. A expressão (1.1.6) significa a integral de a até b da função $\alpha \mapsto \langle\text{fómula}\rangle$; por exemplo:

$$\int_a^b x^2 \,\mathrm{d}x = \frac{1}{3}\,(b^3 - a^3).$$

Note que *não se deve* escrever expressões como

$$\int_a^b f \,\mathrm{d}x,$$

mas pode-se escrever:

$$\int_a^b f(x) \,\mathrm{d}x,$$

[5] Na verdade, é possível também que $b < a$, e nesse caso supomos que o domínio de f contém o intervalo $[b, a]$.

o que é (dado que $f(x)$ é uma fórmula para f) exatamente o mesmo que $\int_a^b f$. A expressão (1.1.6) é em si uma fórmula para uma função que pode ser combinada com $\frac{\mathrm{d}}{\alpha}$ ou com $|_{\alpha=\alpha_0}$, como no exemplo abaixo:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}y}\int_0^{\pi}\cos(x+y)\,\mathrm{d}x\Big|_{y=0} &= \frac{\mathrm{d}}{\mathrm{d}y}\big(\operatorname{sen}(\pi+y)-\operatorname{sen}y\big)\Big|_{y=0}\\ &= \big(\cos(\pi+y)-\cos y\big)\Big|_{y=0} = -2.\end{aligned}$$

Por fim, vamos rever a notação usual para *limites* sob a luz da distinção entre nomes e fórmulas. A notação usual para limites utiliza expressões do tipo:

$$\lim_{\alpha\to\alpha_0}\langle\text{fórmula}\rangle, \tag{1.1.7}$$

onde α deve ser substituído por uma variável e α_0 por uma fórmula. A expressão (1.1.7) significa o limite no ponto α_0 da função $\alpha \mapsto \langle\text{fórmula}\rangle$. Vemos então que a notação usual para limites adere naturalmente a fórmulas de funções; na verdade, não há uma notação padrão para limites que possa ser usada junto ao nome de uma função. No entanto, não há problemas[6]: se f é uma função então $f(x)$ é uma fórmula para f e portanto o limite de f num ponto x_0 pode ser escrito como:

$$\lim_{x\to x_0} f(x).$$

1.1.2. Derivadas parciais e direcionais. Se f é uma função definida num subconjunto (tipicante aberto, mas esse tipo de discussão não vem ao caso aqui) de $\mathbb{R}^n$, $v \in \mathbb{R}^n$ e x pertence ao domínio de f então é comum denotar a derivada direcional de f no ponto x na direção de v por[7]:

$$\frac{\partial f}{\partial v}(x).$$

Essa notação não pode ser usada para fórmulas e portanto utilizaremos:

$$\frac{\partial_\alpha}{\partial v}\langle\text{fórmula}\rangle \tag{1.1.8}$$

onde α pode ser substituído por uma variável qualquer. A expressão (1.1.8) significa $\frac{\partial f}{\partial v}(\alpha)$, onde f é a função $\alpha \mapsto \langle\text{fórmula}\rangle$. Por exemplo, temos:

$$\frac{\partial_x}{\partial v}\operatorname{sen}(x_1^2x_2) = 2\cos(x_1^2x_2)x_1x_2v_1 + \cos(x_1^2x_2)x_1^2v_2, \tag{1.1.9}$$

onde, como é usual, um índice i numa expressão indica que estamos tomando a i-ésima coordenada daquela expressão. Podemos também escrever

[6]Há dificuldade quando temos uma notação que adere a nomes e não temos uma que adere a fórmulas. A situação inversa não causa problemas, já que é simples produzir uma fórmula a partir de um nome, a saber, $f(x)$ é uma fórmula para a função f (obviamente, podemos usar qualquer variável em vez de x).

[7]Alguns textos usam $\frac{\partial f}{\partial \vec{v}}(x)$, mas nós não adotaremos essa notação.

(estendendo um pouco a notação introduzida acima):

$$\frac{\partial_{(x,y)}}{\partial v}(x^2y^3) = 2xy^3v_1 + 3x^2y^2v_2. \tag{1.1.10}$$

Note que em (1.1.9) a variável x representa um vetor[8] de $\mathbb{R}^2$, mas em (1.1.10) a variável x (e também a variável y) representa um número real.

Um caso particular importante das derivadas direcionais são as *derivadas parciais*: se $(e_i)_{i=1}^n$ denota a base canônica de $\mathbb{R}^n$ então a derivada direcional $\frac{\partial f}{\partial e_i}(x)$ de uma função f definida em (ou num subconjunto de) $\mathbb{R}^n$ é também chamada a *i-ésima derivada parcial de f* no ponto x. Na prática, não se usa muito a notação $\frac{\partial f}{\partial e_i}(x)$ para a i-ésima derivada parcial de f no ponto x, mas usa-se, por exemplo, $\frac{\partial f}{\partial x_i}(x)$; além do mais, se f é definida em $\mathbb{R}^3$, muitas vezes se usa $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ e $\frac{\partial f}{\partial z}$ em vez de $\frac{\partial f}{\partial e_1}$, $\frac{\partial f}{\partial e_2}$ e $\frac{\partial f}{\partial e_3}$. Em alguns casos encontra-se também $\frac{\partial f}{\partial u}$, $\frac{\partial f}{\partial v}$, $\frac{\partial f}{\partial \theta}$, etc. O que significa tudo isso? Ocorre que na prática é um tanto desagradável escrever $\frac{\partial f}{\partial e_i}$: digamos que tenhamos uma situação em que $f : \mathbb{R}^{15} \to \mathbb{R}$ modela um certo processo físico que depende de 15 variáveis. Não seria nada prático memorizar os significados físicos de $\frac{\partial f}{\partial e_1}, \ldots, \frac{\partial f}{\partial e_{15}}$. O que se faz então é estabelecer um acordo com o leitor de que os números 1, 2, …, 15 corresponderão a 15 letras que são mais fáceis de lembrar do que os números. Por exemplo, se f é uma função definida em $\mathbb{R}^3$ pode-se combinar que os números 1, 2, 3 corresponderão respectivamente às letras x, y, z, de modo que $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ e $\frac{\partial f}{\partial z}$ devem ser lidos como sinônimos de $\frac{\partial f}{\partial e_1}$, $\frac{\partial f}{\partial e_2}$ e $\frac{\partial f}{\partial e_3}$, respectivamente[9].

1.1.1. OBSERVAÇÃO. Se convencionamos que os números $1, 2, \ldots, n$ correspondem, respectivamente, às letras $x_1, x_2, \ldots, x_n$ então, por definição, $\frac{\partial f}{\partial x_i}(x)$ é o mesmo que $\frac{\partial f}{\partial e_i}(x)$. Observamos que também:

$$\frac{\mathrm{d}}{\mathrm{d}x_i} f(x)$$

é igual a $\frac{\partial f}{\partial e_i}(x)$, mas não apenas por definição. $\frac{\mathrm{d}}{\mathrm{d}x_i} f(x)$ é um caso particular da notação (1.1.4) (entendendo que x abrevia $(x_1, \ldots, x_n)$) e é igual, por definição, a $g'(x_i)$, onde g é a função $x_i \mapsto f(x)$. É fácil ver que $g'(x_i)$ é realmente igual a $\frac{\partial f}{\partial e_i}(x)$, mas por um motivo um pouco mais indireto do que a igualdade entre $\frac{\partial f}{\partial x_i}(x)$ e $\frac{\partial f}{\partial e_i}(x)$, que é resultado de uma mera convenção.

---

[8] Na verdade, o leitor não pode deduzir isso apenas da expressão (1.1.9), já que poderia ser, por exemplo, que $x \in \mathbb{R}^3$, mas estamos considerando uma função que não depende de x_3. O contexto é portanto necessário para a compreensão do significado da expressão.

[9] Um jeito (desnecessariamente) sofisticado de olhar para a situação é o seguinte: em vez de considerar uma função definida em $\mathbb{R}^3$, pensamos numa função definida em $\mathbb{R}^{\{x,y,z\}}$, que é o $\mathbb{R}$-módulo livre gerado por $\{x, y, z\}$. Temos então que x, y e z constituem uma base do espaço vetorial real $\mathbb{R}^{\{x,y,z\}}$. Evidentemente $\mathbb{R}^{\{x,y,z\}}$ é isomorfo a $\mathbb{R}^3$, mas o estabelecimento de um isomorfismo específico depende justamente da escolha de uma bijeção entre as letras x, y, z e os números 1, 2, 3.

1.2. Funções Diferenciáveis

Nesta seção introduzimos o conceito de função diferenciável em espaços vetoriais reais de dimensão finita. Recordamos que todo espaço vetorial real de dimensão finita admite uma norma e que quaisquer duas normas num espaço vetorial real de dimensão finita são *equivalentes*, isto é, determinam os mesmos conjuntos abertos; temos, na verdade, que duas normas equivalentes $\|\cdot\|_1$, $\|\cdot\|_2$ num espaço vetorial real E são também *Lipschitz-equivalentes*, isto é, existem números reais positivos c, c' tais que:

$$c'\|x\|_1 \le \|x\|_2 \le c\|x\|_1, \tag{1.2.1}$$

para todo $x \in E$. O leitor pode consultar a Seção B.11 do Apêndice B para mais detalhes. O fato que todas as normas num espaço vetorial real de dimensão finita são equivalentes implica que podemos lidar com *conceitos topológicos* (tais como continuidade, limites, compacidade, etc) num espaço vetorial real de dimensão finita sem nos preocuparmos em fixar uma norma específica no espaço. O fato de que quaisquer duas normas são Lipschitz-equivalentes nos diz que podemos lidar até mesmo com alguns conceitos que não são puramente topológicos (como completude, continuidade uniforme ou conjunto limitado) sem nos preocuparmos em fixar uma norma específica.

Ao longo de toda a seção consideraremos fixados espaços vetoriais reais de dimensão finita E, F, um subconjunto U de E, uma função $f : U \to F$ e um ponto $x \in U$.

Dada uma aplicação linear $T : E \to F$, nós definimos uma função r fazendo:

$$r(h) = f(x+h) - f(x) - T(h) \in F, \tag{1.2.2}$$

para todo $h \in E$ tal que $x + h \in U$. O domínio da função r é o conjunto:

$$U - x \stackrel{\text{def}}{=} \big\{u - x : u \in U\big\} = \big\{h \in E : x + h \in U\big\},$$

e o contra-domínio de r é o espaço F.

1.2.1. Definição. Suponha que o ponto x não seja isolado em U, isto é, que x seja um ponto de acumulação de U. A função $f : U \to F$ é dita *diferenciável* no ponto x se existe uma aplicação linear $T : E \to F$ tal que a aplicação r definida em (1.2.2) satisfaça a condição:

$$\lim_{h \to 0} \frac{r(h)}{\|h\|} = 0, \tag{1.2.3}$$

onde $\|\cdot\|$ denota uma norma qualquer fixada em E. Mais diretamente, f é diferenciável no ponto x quando existe uma aplicação linear $T : E \to F$ tal que:

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - T(h)}{\|h\|} = 0. \tag{1.2.4}$$

O domínio da função $h \mapsto \frac{r(h)}{\|h\|}$ da qual consideramos o limite em (1.2.3) é o conjunto $(U - x) \setminus \{0\}$; temos que a condição de que x seja um ponto de acumulação de U é necessária e suficiente para que 0 seja um ponto de acumulação de $(U - x) \setminus \{0\}$ e essa última é a condição mínima que precisamos para que o limite em (1.2.3) faça sentido.

Vamos verificar que a condição (1.2.3) não depende das normas fixadas nos espaços E, F. Para isso, observamos primeiro que na condição (1.2.3) há três referências às normas dos espaços E e F: há uma referência à norma de E e uma referência à norma de F escondidas na definição de limite e há uma referência explícita à norma de E na descrição da função da qual tomamos o limite. Com as referências às normas de E e F escondidas na definição de limite não precisamos nos preocupar pois, como mencionado no início da seção, a noção de limite é topológica e todas as normas em espaços vetoriais reais de dimensão finita são equivalentes. Fixemos então nossa atenção na referência a norma de E que aparece no denominador da fração em (1.2.3). Sejam $\|\cdot\|_1$, $\|\cdot\|_2$ normas em E; temos:

$$\frac{r(h)}{\|h\|_1} = \frac{\|h\|_2}{\|h\|_1}\,\frac{r(h)}{\|h\|_2}, \qquad \frac{r(h)}{\|h\|_2} = \frac{\|h\|_1}{\|h\|_2}\,\frac{r(h)}{\|h\|_1},$$

para todo $h \in (U - x) \setminus \{0\}$ e da Lipschitz-equivalência entre $\|\cdot\|_1$ e $\|\cdot\|_2$ (veja (1.2.1)) segue que as funções:

$$E \setminus \{0\} \ni h \longmapsto \frac{\|h\|_2}{\|h\|_1} \in \mathbb{R}, \qquad E \setminus \{0\} \ni h \longmapsto \frac{\|h\|_1}{\|h\|_2} \in \mathbb{R}$$

são limitadas. Daí:

$$\lim_{h \to 0} \frac{r(h)}{\|h\|_1} = 0 \Longleftrightarrow \lim_{h \to 0} \frac{r(h)}{\|h\|_2} = 0.$$

Façamos uma pequena digressão das considerações estritamente formais para explicar melhor a noção de função diferenciável. Temos que no ponto x a função f assume um certo valor $y = f(x)$. Quando o ponto x sofre uma perturbação $\Delta x = h$, o valor y de f no ponto x sofre uma perturbação:

$$\Delta y = f(x + h) - f(x).$$

Num curso de Cálculo para funções de uma variável considera-se o caso em que $E = F = \mathbb{R}$ e a derivada de f no ponto x é definida fazendo o limite:

$$c = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}. \tag{1.2.5}$$

A igualdade (1.2.5) pode ser reescrita na forma:

$$\lim_{\Delta x \to 0} \frac{\Delta y - c\Delta x}{\Delta x} = 0 \tag{1.2.6}$$

e interpretada da seguinte maneira: quando Δx é pequeno, o produto $c\Delta x$ é uma “boa aproximação” para Δy. Por “boa aproximação” entendemos não apenas que $\Delta y - c\Delta x$ tende a zero quando Δx tende a zero[10], mas que $\Delta y - c\Delta x$ tende a zero *mais rápido* do que Δx, no sentido que $\Delta y - c\Delta x$ é o produto de Δx por uma função que tende a zero quando Δx tende a zero (essa função é o quociente que aparece em (1.2.6)). No caso geral (em que E e F não são necessariamente iguais a $\mathbb{R}$) as perturbações Δx e Δy são vetores e o quociente $\frac{\Delta y}{\Delta x}$ sequer faz sentido. No entanto, podemos trocar $c\Delta x$ por $T(\Delta x)$, onde $T : E \to F$ é uma aplicação linear[11] e daí a diferença $\Delta y - T(\Delta x)$ é justamente igual ao *resto* $r(h)$ definido em (1.2.2). Como Δx é um vetor, não faz sentido dividir $\Delta y - T(\Delta x)$ por Δx e a condição de que $\Delta y - T(\Delta x)$ tenda a zero mais rápido do que Δx deve ser expressa através da igualdade:

$$\lim_{\Delta x \to 0} \frac{\Delta y - T(\Delta x)}{\|\Delta x\|} = 0,$$

que é justamente (1.2.3) (note que não faria diferença alguma se em (1.2.6) trocássemos o denominador por $|\Delta x|$, já que isso corresponderia apenas a multiplicar a fração em (1.2.6) pelo sinal de Δx, que é uma função limitada). Pode-se pensar então que f é diferenciável no ponto x quando a perturbação que o valor de f sofre quando perturbamos x pode ser "bem aproximada" por uma função linear da perturbação sofrida por x. Note que *não é correto* dizer (como às vezes se ouve dizer) que f é diferenciável no ponto x quando admite uma "boa aproximação linear" em torno de x: se queremos falar em aproximações de f devemos falar em *aproximações afins* (veja Exercício 1.1).

Já que toda aplicação linear em espaços vetoriais reais de dimensão finita é contínua (veja Proposição B.11.20), é natural esperar que tal propriedade seja herdada pelas funções diferenciáveis. Esse é o conteúdo da seguinte:

1.2.2. Proposição. *Se $f : E \supset U \to F$ é diferenciável no ponto $x \in U$ então f é contínua no ponto x.*

Demonstração. Seja $T : E \to F$ uma aplicação linear satisfazendo (1.2.3), sendo r definida em (1.2.2). Temos:

$$\lim_{h \to 0} f(x+h) = \lim_{h \to 0} \big(f(x) + T(h) + r(h)\big).$$

Já que toda aplicação linear é contínua e $T(0) = 0$ temos $\lim_{h \to 0} T(h) = 0$. Além do mais:

$$\lim_{h \to 0} r(h) = \lim_{h \to 0} \|h\| \frac{r(h)}{\|h\|} = 0,$$

donde $\lim_{h \to 0} f(x+h) = f(x)$ e f é contínua no ponto x. □

[10] A condição de que $\Delta y - c\Delta x$ tende a zero quando Δx tende a zero é equivalente à continuidade de f no ponto x e portanto não diz nada sobre o escalar c.

[11] Se $E = \mathbb{R}^m$ e $F = \mathbb{R}^n$ estamos trocando o escalar c por uma matriz real $n \times m$ (a matriz que representa T com respeito às bases canônicas).

1.2.3. Definição. Dado um vetor $v \in E$ então a *derivada direcional* de f no ponto x e na direção de v é definida por:

$$\frac{\partial f}{\partial v}(x) \overset{\text{def}}{=} \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t} \in F, \tag{1.2.7}$$

desde que 0 seja um ponto de acumulação do conjunto:

$$\big\{t \in \mathbb{R} : x + tv \in U\big\} \tag{1.2.8}$$

e que o limite que aparece do lado direito da igualdade em (1.2.7) exista.

O domínio da função $t \mapsto \frac{f(x+tv)-f(x)}{t}$ da qual consideramos o limite em (1.2.7) é o conjunto:

$$\big\{t \in \mathbb{R} : x + tv \in U\big\} \setminus \{0\}; \tag{1.2.9}$$

temos que 0 é ponto de acumulação de (1.2.8) se e somente se 0 é ponto de acumulação de (1.2.9) e essa última é a condição mínima necessária para que faça sentido considerar o limite em (1.2.7).

1.2.4. Observação. Se $v = 0$ então o conjunto (1.2.8) é $\mathbb{R}$ e a derivada direcional $\frac{\partial f}{\partial v}(x)$ certamente existe e é igual a zero. Se $v \neq 0$ então a aplicação $\mathbb{R} \ni t \mapsto x + tv$ é um homeomorfismo sobre a reta:

$$x + \mathbb{R}v = \big\{x + tv : t \in \mathbb{R}\big\}$$

e 0 é ponto de acumulação do conjunto (1.2.8) se e somente se x é ponto de acumulação do conjunto $(x + \mathbb{R}v) \cap U$; vemos que se $v \neq 0$ então para que 0 seja ponto de acumulação de (1.2.8) é necessário (mas em geral longe de ser suficiente) que x seja ponto de acumulação de U. No caso particular em que E tem dimensão 1, a reta $x + \mathbb{R}v$ coincide com E para $v \neq 0$ de modo que 0 é ponto de acumulação de (1.2.8) se e somente se x é ponto de acumulação de U.

1.2.5. Observação. Alguns livros elementares de Cálculo definem a noção de derivada direcional apenas quando o vetor direção v tem norma igual a 1. Não há realmente nenhuma justificativa matemática para se fazer isso e, na verdade, será bastante conveniente para nós permitir derivadas direcionais $\frac{\partial f}{\partial v}(x)$ com $v \in E$ arbitrário (veja, por exemplo, Lema 1.2.8 a seguir). Alguns livros adotam a convenção de definir $\frac{\partial f}{\partial v}(x)$ apenas quando $\|v\| = 1$ pois quando se está interessado em estudar a forma como f varia ao longo de uma reta que passa por um ponto x faz sentido eleger um vetor unitário paralelo a essa reta como uma espécie de "representante natural" da direção determinada por essa reta. No Exercício 1.2 pedimos ao leitor para relacionar as derivadas direcionais $\frac{\partial f}{\partial v}(x)$ e $\frac{\partial f}{\partial (cv)}(x)$, sendo $c \in \mathbb{R}$ um escalar.

1.2.6. Definição. Se $E = \mathbb{R}$ e $x \in U$ é um ponto de acumulação de $U \subset \mathbb{R}$ então a derivada direcional:

$$\frac{\partial f}{\partial 1}(x) = \lim_{t \to 0} \frac{f(x + t) - f(x)}{t} \in F$$

é denotada também por $f'(x)$ e é chamada a *derivada* (ou o *vetor tangente*) de f no ponto (ou *no instante*) x.

Quando $E = \mathbb{R}$ podemos pensar f como sendo uma curva parametrizada no espaço vetorial F, o que explica a adoção da terminologia "vetor tangente". Quando F também é igual a $\mathbb{R}$ a Definição 1.2.6 coincide com a definição usual de derivada dos cursos de Cálculo de uma variável.

1.2.7. Exercício. Seja $v \in E$ e suponha que 0 é um ponto de acumulação do conjunto (1.2.8). Considere a função:

$$\phi : \mathbb{R} \supset \big\{t \in \mathbb{R} : x + tv \in U\big\} \ni t \longmapsto f(x + tv) \in F.$$

Mostre que a derivada direcional $\frac{\partial f}{\partial v}(x)$ existe se e somente se a derivada $\phi'(0)$ existe e que, caso ambas existam, são iguais.

No Exercício 1.3 pedimos ao leitor para demonstrar uma pequena generalização do resultado do Exercício 1.2.7.

Vamos agora relacionar a noção de função diferenciável com a noção de derivada direcional.

1.2.8. Lema. *Suponha que $f : E \supset U \to F$ seja diferenciável no ponto $x \in U$ e que $T : E \to F$ seja uma aplicação linear tal que a condição* (1.2.4) *seja satisfeita. Se $v \in E$ é tal que 0 é ponto de acumulação do conjunto* (1.2.8) *então a derivada direcional $\frac{\partial f}{\partial v}(x)$ existe e é igual a $T(v)$.*

Demonstração. Se $v = 0$ então $\frac{\partial f}{\partial v}(x) = 0 = T(v)$. Se $v \neq 0$ podemos fazer $h = tv$ em (1.2.4) para obter:

$$\lim_{t \to 0} \frac{f(x + tv) - f(x) - T(tv)}{|t|\|v\|} = 0. \tag{1.2.10}$$

Como:

$$\frac{f(x + tv) - f(x) - T(tv)}{t} = \|v\| \, \frac{|t|}{t} \, \frac{f(x + tv) - f(x) - T(tv)}{|t|\|v\|}$$

e a função $t \mapsto \frac{|t|}{t}$ é limitada concluímos que:

$$\lim_{t \to 0} \frac{f(x + tv) - f(x) - T(tv)}{t} = 0.$$

Mas:

$$\frac{f(x + tv) - f(x) - T(tv)}{t} = \frac{f(x + tv) - f(x)}{t} - T(v),$$

donde segue que:

$$\lim_{t \to 0} \frac{f(x + tv) - f(x)}{t} = T(v). \qquad \square$$

1.2.9. Corolário. *Se x é um ponto interior de U e $f : E \supset U \to F$ é diferenciável no ponto x então existe apenas uma transformação linear $T : E \to F$ tal que a condição* (1.2.4) *é satisfeita. Além do mais, para todo $v \in E$ a derivada direcional $\frac{\partial f}{\partial v}(x)$ existe e é igual a $T(v)$.*

Demonstração. Já que x é um ponto interior de U temos que para todo $v \in E$ o conjunto (1.2.8) contém uma vizinhança de 0 e em particular possui 0 como ponto de acumulação. Se $T : E \to F$ é uma aplicação linear satisfazendo (1.2.4) podemos então aplicar o Lema 1.2.4 para obter:

$$T(v) = \frac{\partial f}{\partial v}(x),$$

para todo $v \in E$. Essa igualdade evidentemente prova a unicidade de T. $\square$

1.2.10. Corolário. *Se E tem dimensão* 1*, $x \in U$ é um ponto de acumulação de U e $f : E \supset U \to F$ é diferenciável no ponto x então existe apenas uma transformação linear $T : E \to F$ tal que a condição* (1.2.4) *seja satisfeita.*

Demonstração. Como E tem dimensão 1 e x é ponto de acumulação de U, temos que 0 é ponto de acumulação do conjunto (1.2.8) para todo $v \in E$ (veja Observação 1.2.4). Podemos então argumentar como na demonstração do Corolário 1.2.9. $\square$

Os Corolários 1.2.9 e 1.2.10 nos permitem enunciar a seguinte:

1.2.11. Definição. Se x é um ponto interior de U ou se E tem dimensão 1 e $x \in U$ é um ponto de acumulação de U e se $f : E \supset U \to F$ é diferenciável no ponto x então a única aplicação linear $T : E \to F$ tal que a condição (1.2.4) é satisfeita é chamada a *diferencial* de f no ponto x e é denotada por $\mathrm{d}f(x)$.

Evidentemente, se x é um ponto interior de U ou se E tem dimensão 1 e $x \in U$ é um ponto de acumulação de U e se f é diferenciável no ponto x então:

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - \mathrm{d}f(x) \cdot h}{\|h\|} = 0. \tag{1.2.11}$$

Alguns esclarecimentos são necessários sobre a notação $\mathrm{d}f(x) \cdot h$: temos que a diferencial $\mathrm{d}f(x)$ é uma função (uma aplicação linear de E para F) e portanto pode ser aplicada a um vetor $h \in E$. A notação padrão para aplicar uma função a um ponto de seu domínio nos faria escrever $\mathrm{d}f(x)(h)$ para denotar o valor de $\mathrm{d}f(x)$ no ponto h. No entanto, para evitar sobrecarga de parênteses, nós muitas vezes escreveremos apenas $\mathrm{d}f(x) \cdot h$ (o símbolo $\cdot$ pode ser lido como "aplicado em") ou apenas $\mathrm{d}f(x)h$.

Dado um vetor $v \in E$, temos que $\mathrm{d}f(x) \cdot v$ coincide com a derivada direcional de f no ponto x e na direção de v. De fato, o Lema 1.2.8 implica diretamente a seguinte:

1.2.12. Proposição. *Se x é um ponto interior de U ou se E tem dimensão* 1 *e $x \in U$ é um ponto de acumulação de U e se $f : E \supset U \to F$ é diferenciável no ponto x então para todo $v \in E$ a derivada direcional $\frac{\partial f}{\partial v}(x)$*

existe e é igual a $\mathrm{d}f(x) \cdot v$*, isto é:*

$$\frac{\partial f}{\partial v}(x) = \mathrm{d}f(x) \cdot v,$$

para todo $v \in E$. □

Será possível definir a direrencial $\mathrm{d}f(x)$ em condições mais gerais do que aquelas enunciadas na Definição 1.2.11? Em outras palavras, será possível mostrar a unicidade da aplicação linear $T : E \to F$ satisfazendo (1.2.4) em condições mais gerais do que aquelas que aparecem nos Corolários 1.2.9 e 1.2.10? A resposta é sim: para que possamos provar a unicidade de T (e então definir $\mathrm{d}f(x)$) precisamos apenas que x seja ponto de acumulação do domínio U "ao longo de uma quantidade de direções grande o suficiente para gerar o espaço E". Pedimos ao leitor para trabalhar os detalhes dessa questão nos Exercícios 1.4, 1.5 e 1.6. No entanto, para simplificar a exposição, nós continuaremos a trabalhar com a noção de diferencial apenas no caso em que o ponto x é interior a U ou que E tem dimensão 1 e $x \in U$ é um ponto de acumulação de U.

Vejamos alguns exemplos bem simples de aplicações diferenciáveis e suas respectivas diferenciais.

1.2.13. Exemplo. Se $f : E \to F$ é constante então f é diferenciável em qualquer ponto $x \in E$ e $\mathrm{d}f(x) = 0$. De fato, se $T = 0$ então a função r definida em (1.2.2) é nula.

1.2.14. Exemplo. Se $f : E \to F$ é linear então para todo $x \in E$ a função f é diferenciável no ponto x e $\mathrm{d}f(x) = f$. De fato, fazendo $T = f$ temos que a aplicação r definida em (1.2.2) é nula.

1.2.15. Exercício. Seja V um subconjunto de U e suponha que x é um ponto interior de V ou que E tem dimensão 1 e $x \in V$ é um ponto de acumulação de V. Mostre que:

(a) se $f : E \supset U \to F$ é diferenciável no ponto x então também $f|_V : V \to F$ é diferenciável no ponto x e $\mathrm{d}(f|_V)(x) = \mathrm{d}f(x)$;

(b) se V é uma vizinhança de x relativamente a U e se $f|_V$ é diferenciável no ponto x então também f é diferenciável no ponto x.

A existência de uma quantidade grande de derivadas direcionais de f no ponto x implica na diferenciabilidade de f no ponto x? Não. É possível que x seja um ponto interior de U e que todas as derivadas direcionais de f no ponto x existam, mas que f não seja diferenciável no ponto x (os resultados dos Exercícios 1.7 e 1.8 fornecem uma família de exemplos que confirmam essa afirmação). No entanto, quando E tem dimensão 1, temos o seguinte:

1.2.16. Lema. *Se* E *tem dimensão* 1*,* $x \in U$ *é um ponto de acumulação de* U *e se a derivada direcional* $\frac{\partial f}{\partial v}(x)$ *existe para algum* $v \in E$ *não nulo então* f *é diferenciável no ponto* x*.*

Demonstração. Como E tem dimensão 1 e v é não nulo temos que o conjunto unitário $\{v\}$ é uma base de E e portanto existe uma única aplicação linear $T : E \to F$ tal que $T(v) = \frac{\partial f}{\partial v}(x)$. Vamos mostrar que a condição (1.2.4) é satisfeita. Como a aplicação $\mathbb{R} \ni t \mapsto tv \in E$ é um homeomorfismo, temos que (1.2.4) é equivalente a (1.2.10). Mas:

$$\frac{f(x+tv) - f(x) - T(tv)}{|t|\|v\|} = \frac{1}{\|v\|}\,\frac{t}{|t|}\Big(\frac{f(x+tv) - f(x)}{t} - T(v)\Big),$$

onde a função $t \mapsto \frac{t}{|t|}$ é limitada e:

$$\lim_{t\to 0}\Big(\frac{f(x+tv) - f(x)}{t} - T(v)\Big) = \frac{\partial f}{\partial v}(x) - T(v) = 0,$$

donde (1.2.10) vale. □

1.2.17. Corolário. *Se $E = \mathbb{R}$ e x é um ponto de acumulação de $U \subset \mathbb{R}$ então a função $f : U \to F$ é diferenciável no ponto x se e somente se a derivada $f'(x)$ existe (veja Definição 1.2.6). Caso f seja diferenciável no ponto x, temos:*

$$\mathrm{d}f(x)\cdot h = f'(x)h, \tag{1.2.12}$$

para todo $h \in \mathbb{R}$ e $f'(x) = \mathrm{d}f(x)\cdot 1$.

Demonstração. Se f é diferenciável no ponto x então segue da Proposição 1.2.12 que a derivada direcional $\frac{\partial f}{\partial 1}(x) = f'(x)$ existe e é igual a $\mathrm{d}f(x)\cdot 1$; da linearidade de $\mathrm{d}f(x)$ vem a igualdade (1.2.12). Reciprocamente, se $f'(x) = \frac{\partial f}{\partial 1}(x)$ existe então o Lema 1.2.16 implica que f é diferenciável no ponto x. □

1.2.18. Convenção. Em tudo que segue, assumimos que x seja um ponto interior de U ou que E tenha dimensão 1 e $x \in U$ seja um ponto de acumulação de U.

1.2.19. Proposição. *Se $f : E \supset U \to F$, $g : E \supset U \to F$ são ambas aplicações diferenciáveis no ponto x então a aplicação $f + g$ também é diferenciável no ponto x e:*

$$\mathrm{d}(f+g)(x) = \mathrm{d}f(x) + \mathrm{d}g(x).$$

Demonstração. Basta notar que $\mathrm{d}f(x) + \mathrm{d}g(x)$ é uma aplicação linear e que:

$$\begin{aligned}\lim_{h\to 0}\frac{(f+g)(x+h) - (f+g)(x) - \big(\mathrm{d}f(x) + \mathrm{d}g(x)\big)(h)}{\|h\|}\\ = \lim_{h\to 0}\frac{f(x+h) - f(x) - \mathrm{d}f(x)\cdot h}{\|h\|}\\ + \lim_{h\to 0}\frac{g(x+h) - g(x) - \mathrm{d}g(x)\cdot h}{\|h\|} = 0.\end{aligned}$$ □

1.2.20. Lema. *Seja F' um espaço vetorial real de dimensão finita. Se $f : E \supset U \to F$ é diferenciável no ponto x e se $L : F \to F'$ é uma aplicação linear então $L \circ f$ também é diferenciável no ponto x e:*

$$\mathrm{d}(L \circ f)(x) = L \circ \mathrm{d}f(x).$$

Demonstração. Como L é contínua e $L(0) = 0$, de (1.2.11) vem:

$$\lim_{h \to 0} L\Big(\frac{f(x+h) - f(x) - \mathrm{d}f(x) \cdot h}{\|h\|}\Big) = 0;$$

daí:

$$\lim_{h \to 0} \frac{(L \circ f)(x+h) - (L \circ f)(x) - \big(L \circ \mathrm{d}f(x)\big)(h)}{\|h\|} = 0,$$

e a conclusão segue. □

1.2.21. Corolário. *Se $f : E \supset U \to F$ é diferenciável no ponto x e se $c \in \mathbb{R}$ então a aplicação:*

$$cf : U \ni y \longmapsto cf(y) \in F$$

é diferenciável no ponto x e:

$$\mathrm{d}(cf)(x) = c\big(\mathrm{d}f(x)\big) : E \ni h \longmapsto c\big(\mathrm{d}f(x) \cdot h\big) \in F.$$

Demonstração. Aplique o Lema 1.2.20 com $L : F \ni w \mapsto cw \in F$. □

1.2.22. Corolário. *Se $\phi : E \supset U \to \mathbb{R}$ é diferenciável no ponto x e se $w \in F$ então a aplicação:*

$$\phi w : U \ni y \longmapsto \phi(y)w \in F$$

é diferenciável no ponto x e:

$$\mathrm{d}(\phi w)(x) = \big(\mathrm{d}\phi(x)\big)w : E \ni h \longmapsto \big(\mathrm{d}\phi(x) \cdot h\big)w \in F.$$

Demonstração. Aplique o Lema 1.2.20 com $L : \mathbb{R} \ni c \mapsto cw \in F$. □

1.2.23. Corolário. *Suponha que $\mathcal{B} = (e_i)_{i=1}^n$ seja uma base de F e que $\pi_i : F \to \mathbb{R}$ denote o funcional linear que associa a cada vetor de F sua i-ésima coordenada na base $\mathcal{B}$. Seja $f_i = \pi_i \circ f : U \to \mathbb{R}$, $i = 1, \ldots, n$. Temos que f é diferenciável no ponto x se e somente se f_i é diferenciável no ponto x para todo $i = 1, \ldots, n$; além do mais, se f é diferenciável no ponto x então para todo $h \in E$, $\mathrm{d}f_i(x) \cdot h$ é a i-ésima coordenada de $\mathrm{d}f(x) \cdot h$ na base $\mathcal{B}$, isto é:*

$$\mathrm{d}f(x) \cdot h = \sum_{i=1}^{n} \big(\mathrm{d}f_i(x) \cdot h\big)e_i.$$

Demonstração. Se f é diferenciável no ponto x então segue diretamente do Lema 1.2.20 que $f_i = \pi_i \circ f$ é diferenciável no ponto x, já que π_i é linear; além do mais, para todo $h \in E$, $\mathrm{d}f_i(x) \cdot h = \pi_i\big(\mathrm{d}f(x) \cdot h\big)$, isto é, $\mathrm{d}f_i(x) \cdot h$ é a i-ésima coordenada de $\mathrm{d}f(x) \cdot h$ na base $\mathcal{B}$. Reciprocamente, se f_i é diferenciável no ponto x para todo $i = 1, \ldots, n$ então, como

$f = \sum_{i=1}^{n} f_i e_i$, segue da Proposição 1.2.19 e do Corolário 1.2.22 que f é diferenciável no ponto x. $\square$

1.2.24. Corolário. *Suponha que F' seja um subespaço de F e que a imagem de $f : E \supset U \to F$ esteja contida em F'; denote por $f_0 : U \to F'$ a aplicação que difere de f apenas pelo contra-domínio. Temos que f é diferenciável no ponto x se e somente se f_0 é diferenciável no ponto x; além do mais, se f é diferenciável no ponto x então a imagem de $\mathrm{d}f(x) : E \to F$ está contida em F' e as aplicações lineares $\mathrm{d}f(x) : E \to F$ e $\mathrm{d}f_0(x) : E \to F'$ diferem apenas pelo contra-domínio.*

Demonstração. Denote por $i : F' \to F$ a aplicação inclusão. Já que $f = i \circ f_0$ e i é linear, o Lema 1.2.20 nos diz que se f_0 é diferenciável no ponto x então f também é diferenciável no ponto x e $\mathrm{d}f(x) = i \circ \mathrm{d}f_0(x)$, isto é, $\mathrm{d}f(x)$ toma valores em F' e as aplicações $\mathrm{d}f(x)$ e $\mathrm{d}f_0(x)$ diferem apenas pelo contra-domínio. Reciprocamente, suponha que f seja diferenciável no ponto x. Seja $L : F \to F'$ uma aplicação linear cuja restrição a F' seja a aplicação identidade de F' (para obter L, escolha um subespaço complementar qualquer de F' em F e tome L como sendo a projeção em F' relativamente à decomposição em soma direta obtida). Temos então que $f_0 = L \circ f$ e portanto o Lema 1.2.20 nos dá que f_0 é diferenciável no ponto x. $\square$

Exercícios para o Capítulo 1

Diferenciação.

1.1. Exercício. Sejam dados espaços vetoriais E, F. Uma aplicação $A : E \to F$ é dita *afim* se existem uma aplicação linear $T : E \to F$ e um vetor $u \in F$ tais que $A(z) = u + T(z)$, para todo $z \in E$. Suponha que E, F sejam espaços vetoriais reais de dimensão finita e que $f : U \to F$ seja uma função definida num subconjunto U de E. Seja $x \in U$ um ponto de acumulação de U. Dizemos que uma aplicação afim $A : E \to F$ é uma *aproximação de primeira ordem* para f no ponto x se $A(x) = f(x)$ e se:

$$\lim_{z \to x} \frac{f(z) - A(z)}{\|z - x\|} = 0,$$

onde $\|\cdot\|$ denota uma norma qualquer fixada em E. Mostre que:

(a) se f é diferenciável no ponto x e $T : E \to F$ é uma aplicação linear que satisfaz (1.2.4) então a aplicação afim:

$$A : E \ni z \longmapsto f(x) - T(x) + T(z) \in F$$

é uma aproximação de primeira ordem para f;

(b) se existe uma aplicação afim $A : E \ni z \mapsto u + T(z) \in F$ (com $T : E \to F$ linear e $u \in F$) que é uma aproximação de primeira ordem para f então $u = f(x) - T(x)$, T satisfaz (1.2.4) e portanto f é diferenciável no ponto x.

1.2. Exercício. Sejam E, F espaços vetoriais reais de dimensão finita, U um subconjunto de E, $x \in U$, $v \in E$ e $f : U \to F$ uma função. Suponha que 0 seja um ponto de acumulação do conjunto $\big\{t \in \mathbb{R} : x + tv \in U\big\}$ e que a derivada direcional $\frac{\partial f}{\partial v}(x)$ exista. Dado $c \in \mathbb{R}$, mostre que 0 é ponto de acumulação do conjunto $\big\{t \in \mathbb{R} : x + tcv \in U\big\}$, que a derivada direcional $\frac{\partial f}{\partial (cv)}(x)$ existe e que:

$$\frac{\partial f}{\partial (cv)}(x) = c\,\frac{\partial f}{\partial v}(x).$$

1.3. Exercício. Sejam E, F espaços vetoriais reais de dimensão finita, U um subconjunto de E, $x \in U$, $v \in E$ e $f : U \to F$ uma função. Considere a função:

$$\phi : \mathbb{R} \supset \big\{t \in \mathbb{R} : x + tv \in U\big\} \ni t \longmapsto f(x + tv) \in F.$$

Dado $t_0 \in \mathbb{R}$, mostre que:

(a) t_0 é ponto de acumulação do domínio de ϕ se e somente se 0 é ponto de acumulação do conjunto $\big\{t \in \mathbb{R} : (x + t_0 v) + tv \in U\big\}$;

(b) se t_0 pertence ao domínio de ϕ e é ponto de acumulação do domínio de ϕ então a derivada $\phi'(t_0)$ existe se e somente se a derivada direcional $\frac{\partial f}{\partial v}(x + t_0 v)$ existe e, caso ambas existam, são iguais.

1.4. Exercício. Sejam E um espaço vetorial real de dimensão finita, U um subconjunto de E e $x, v \in E$. Dizemos que x é um *ponto de acumulação de U na direção de v* se existe uma seqüência $(x_n)_{n\geq 0}$ em U e uma seqüência $(t_n)_{n\geq 0}$ de números reais positivos tal que $x_n \to x$ e $t_n(x_n - x) \to v$. Se $\|\cdot\|$ é uma norma em E e $v \neq 0$, mostre que x é um ponto de acumulação de U na direção de v se e somente se existe uma seqüência $(x_n)_{n\geq 0}$ em $U \setminus \{x\}$ tal que $x_n \to x$ e:

$$\frac{x_n - x}{\|x_n - x\|} \longrightarrow \frac{v}{\|v\|}.$$

1.5. Exercício. Sejam E, F espaços vetoriais reais de dimensão finita, U um subconjunto de E, $f : U \to F$ uma função, $x \in U$ e $v \in E$ um vetor tal que x é ponto de acumulação de U na direção de v. Se $T : E \to F$, $T' : E \to F$ são aplicações lineares tais que:

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - T(h)}{\|h\|} = 0, \tag{1.13}$$

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - T'(h)}{\|h\|} = 0,$$

mostre que $T(v) = T'(v)$. Conclua que se f é diferenciável no ponto x e o conjunto:

$$\big\{v \in E : x \text{ é ponto de acumulação de } U \text{ na direção de } v\big\} \tag{1.14}$$

gera E como espaço vetorial então existe uma única aplicação linear T tal que a condição (1.13) é satisfeita.

1.6. Exercício. Sejam E, F espaços vetoriais reais de dimensão finita, U um subconjunto de E e $x \in U$. Denote por $f : U \to F$ a função identicamente nula. Mostre que para qualquer aplicação linear $T : E \to F$ que se anula no conjunto (1.14) a condição (1.13) é satisfeita. Conclua que se (1.14) não gera E como espaço vetorial então existem infinitas aplicações lineares T tais que a condição (1.13) é satisfeita.

1.7. Exercício. Sejam E, F espaços vetoriais reais de dimensão finita e denote por S a esfera unitária:

$$S = \big\{x \in E : \|x\| = 1\big\},$$

onde $\|\cdot\|$ é uma norma arbitrária fixada em E. Seja $g : S \to F$ uma função e defina $f : E \to F$ fazendo $f(x) = \|x\|\, g\big(\frac{x}{\|x\|}\big)$, para todo $x \in E \setminus \{0\}$ e $f(0) = 0$. Mostre que:

(a) dado $v \in S$ então a derivada direcional $\frac{\partial f}{\partial v}(0)$ existe se e somente se $g(-v) = -g(v)$ e, em caso afirmativo, essa derivada direcional é igual a $g(v)$;
(b) f é contínua no ponto 0 se e somente se a função g é limitada;
(c) f é diferenciável no ponto 0 se e somente se existe uma aplicação linear $T : E \to F$ tal que $g = T|_S$ e, em caso afirmativo, $\mathrm{d}f(0) = T$.

1.8. Exercício. Sejam E, S como no enunciado do Exercício 1.7 e seja $\phi : S \to \,]0, +\infty[$ uma função. Defina $f : E \to \mathbb{R}$ fazendo $f(0) = 0$, $f(x) = 0$ se $x \neq 0$ e $\|x\| < \phi\big(\frac{x}{\|x\|}\big)$, e $f(x) = 1$ se $x \neq 0$ e $\|x\| \geq \phi\big(\frac{x}{\|x\|}\big)$. Mostre que:

(a) para todo $v \in E$ a derivada direcional $\frac{\partial f}{\partial v}(0)$ existe e é igual a zero;
(b) f é contínua no ponto 0 se e somente se $\inf \phi(S) > 0$;
(c) f é diferenciável no ponto 0 se e somente se $\inf \phi(S) > 0$.

APÊNDICE A

Um pouco de álgebra linear e multilinear

Todos os espaços vetoriais que aparecem neste apêndice são espaços vetoriais sobre um certo corpo de escalares $\mathbb{K}$ que é considerado fixado ao longo de todo o apêndice. O leitor pode, se desejar, assumir que esse corpo é $\mathbb{R}$ ou $\mathbb{C}$ (que são os únicos corpos que são relevantes nestas notas), mas na verdade nenhuma particularidade desses corpos é usada nas demonstrações[1].

A.1. Aplicações lineares e multilineares

Se E, F são espaços vetoriais então uma aplicação $T : E \to F$ é dita *linear* se $T(x + y) = T(x) + T(y)$ e se $T(\lambda x) = \lambda T(x)$, para todos $x, y \in E$ e para todo $\lambda \in \mathbb{K}$. Dados espaços vetoriais E_1, E_2, F então uma aplicação $B : E_1 \times E_2 \to F$ é dita *bilinear* se for "linear em cada uma de suas variáveis", isto é, se para todo $x \in E_1$ a aplicação $E_2 \ni y \mapsto B(x, y) \in F$ é linear e para todo $y \in E_2$ a aplicação $E_1 \ni x \mapsto B(x, y) \in F$ é linear; em outras palavras, devemos ter:

$$\begin{gathered} B(x + x', y) = B(x, y) + B(x', y), \\ B(x, y + y') = B(x, y) + B(x, y'), \\ B(\lambda x, y) = \lambda B(x, y), \quad B(x, \lambda y) = \lambda B(x, y) \end{gathered} \tag{A.1.1}$$

para todos $x, x' \in E_1$, $y, y' \in E_2$ e para todo $\lambda \in \mathbb{K}$.

A.1.1. Exercício. Dados espaços vetoriais E_1, E_2, F e uma aplicação bilinear $B : E_1 \times E_2 \to F$, mostre que:

$$B\Big(\sum_{i=1}^{n} x_i, y\Big) = \sum_{i=1}^{n} B(x_i, y), \quad B\Big(x, \sum_{i=1}^{n} y_i\Big) = \sum_{i=1}^{n} B(x, y_i),$$

para todos $x_1, \ldots, x_n, x \in E_1$, $y_1, \ldots, y_n, y \in E_2$.

A.1.2. Notação. $\mathrm{Lin}(E, F)$ denota o conjunto de todas as aplicações lineares $T : E \to F$ e $\mathrm{Lin}(E_1, E_2; F)$ denota o conjunto de todas as aplicações bilineares $B : E_1 \times E_2 \to F$, onde E, E_1, E_2, F são espaços vetoriais.

A.1.3. Exercício. Dados espaços vetoriais E, F, mostre que $\mathrm{Lin}(E, F)$ é um espaço vetorial quando munido das operações definidas pelas igualdades:

$$(T + T')(x) = T(x) + T'(x), \quad (\lambda T)(x) = \lambda\big(T(x)\big),$$

[1]Na verdade, tudo que aparece neste apêndice poderia ser feito com módulos sobre um anel comutativo arbitrário; a comutatividade do anel de escalares é necessária quando tratamos de aplicações multilineares.

onde $x \in E$, $\lambda \in \mathbb{K}$ e $T, T' \in \operatorname{Lin}(E, F)$. Similarmente, dados espaços vetoriais E_1, E_2, mostre que $\operatorname{Lin}(E_1, E_2; F)$ é um espaço vetorial quando munido das operações definidas pelas igualdades:

$$(B + B')(x, y) = B(x, y) + B'(x, y), \quad (\lambda B)(x, y) = \lambda\big(B(x, y)\big),$$

onde $x \in E_1$, $y \in E_2$, $\lambda \in \mathbb{K}$ e $B, B' \in \operatorname{Lin}(E_1, E_2; F)$.

A.1.4. Exemplo. O *produto interno canônico* de $\mathbb{R}^n$:

$$\mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \longmapsto \langle x, y\rangle = \sum_{i=1}^{n} x_i y_i \in \mathbb{R}$$

é uma aplicação bilinear. Mais geralmente, qualquer produto interno num espaço vetorial real é (por definição) uma aplicação bilinear (veja o enunciado do Exercício A.1 para a definição geral de produto interno em espaços vetoriais reais).

A.1.5. Exemplo. O *produto vetorial* em $\mathbb{R}^3$:

$$\mathbb{R}^3 \ni (x, y) \longmapsto x \times y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1) \in \mathbb{R}^3$$

é uma aplicação bilinear.

A.1.6. Notação. $\mathrm{M}_{m\times n}(\mathbb{K})$ denota o espaço vetorial das matrizes $m \times n$ com entradas em $\mathbb{K}$, onde m, n são números naturais. Escrevemos também $\mathrm{M}_n(\mathbb{K}) = \mathrm{M}_{n\times n}(\mathbb{K})$.

A.1.7. Exemplo. A multiplicação de números complexos é uma aplicação bilinear de $\mathbb{C} \times \mathbb{C}$ em $\mathbb{C}$ (aqui podemos pensar em $\mathbb{C}$ como um espaço vetorial sobre $\mathbb{R}$ ou como um espaço vetorial sobre $\mathbb{C}$)[2].

A.1.8. Exemplo. Dados números naturais m, n, p então a multiplicação de matrizes:

$$\mathrm{M}_{m\times n}(\mathbb{K}) \times \mathrm{M}_{n\times p}(\mathbb{K}) \ni (A, B) \longmapsto AB \in \mathrm{M}_{m\times p}(\mathbb{K})$$

é uma aplicação bilinear.

A.1.9. Exemplo. Dados espaços vetoriais E, F, então a *aplicação de avaliação* definida por:

$$\mathrm{aval} : \operatorname{Lin}(E, F) \times E \ni (T, x) \longmapsto T(x) \in F$$

é bilinear. Dado um outro espaço vetorial G então a *aplicação de composição de aplicações lineares*:

$$\operatorname{Lin}(F, G) \times \operatorname{Lin}(E, F) \ni (S, T) \longmapsto S \circ T \in \operatorname{Lin}(E, G)$$

é bilinear.

[2] A multiplicação de quatérnios também define uma aplicação bilinear se pensarmos no conjunto dos quatérnios como um espaço vetorial sobre $\mathbb{R}$, *mas não se pensarmos nos quatérnios como um espaço vetorial sobre* $\mathbb{C}$. Observamos que, mais geralmente, se $\mathfrak{A}$ é uma álgebra sobre $\mathbb{K}$ então a multiplicação de $\mathfrak{A}$ é (por definição) uma aplicação bilinear de $\mathfrak{A} \times \mathfrak{A}$ em $\mathfrak{A}$.

Se $T : E \to F$, $T' : E \to F$ são aplicações lineares e se $C \subset E$ é um conjunto de geradores para E (isto é, se todo elemento de E é uma combinação linear finita de elementos de C ou, equivalentemente, se nenhum subespaço próprio de E contém C) e se T, T' coincidem em elementos de C então $T = T'$. Algo similar vale para aplicações bilineares:

A.1.10. PROPOSIÇÃO. *Sejam E_1, E_2, F espaços vetoriais, C_1, C_2 conjuntos de geradores para E_1, E_2, respectivamente, e $B, B' \in \mathrm{Lin}(E_1, E_2; F)$ aplicações bilineares. Se $B(x, y) = B'(x, y)$ para todos $x \in C_1$, $y \in C_2$ então $B = B'$.*

DEMONSTRAÇÃO. Dados $x \in E_1$, $y \in E_2$, podemos escrever:

$$x = \sum_{i=1}^{n} \lambda_i x_i, \quad y = \sum_{j=1}^{m} \mu_j y_j,$$

com $x_1, \dots, x_n \in C_1$, $y_1, \dots, y_m \in C_2$ e $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_m \in \mathbb{K}$. Daí:

$$\begin{aligned} B(x, y) = B\Big(\sum_{i=1}^{n} \lambda_i x_i, \sum_{j=1}^{m} \mu_j y_j\Big) &= \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_i \mu_j B(x_i, y_j) \\ = \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_i \mu_j B'(x_i, y_j) &= B'\Big(\sum_{i=1}^{n} \lambda_i x_i, \sum_{j=1}^{m} \mu_j y_j\Big) = B'(x, y). \end{aligned}$$

Logo $B = B'$. □

Dados uma base $\mathcal{B} = (e_i)_{i=1}^n$ de um espaço vetorial E e elementos $(f_i)_{i=1}^n$ de um espaço vetorial F então existe uma única aplicação linear $T : E \to F$ tal que $T(e_i) = f_i$, para $i = 1, \dots, n$. Essa aplicação linear T é definida por:

$$T(x) = \sum_{i=1}^{n} \pi_i(x) f_i,$$

onde $\pi_i : E \to \mathbb{K}$ é o funcional linear que associa a cada $x \in E$ a sua i-ésima coordenada na base $\mathcal{B}$ (recorde que $(\pi_i)_{i=1}^n$ é uma base do *espaço dual* $E^* = \mathrm{Lin}(E, \mathbb{K})$ que é normalmente chamada a *base dual* de $\mathcal{B}$). Um resultado similar vale para aplicações bilineares:

A.1.11. PROPOSIÇÃO. *Sejam E_1, E_2, F espaços vetoriais, $\mathcal{B}^1 = (e_i^1)_{i=1}^m$ uma base de E_1, $\mathcal{B}^2 = (e_i^2)_{i=1}^n$ uma base de E_2 e $f_{ij} \in F$, $i = 1, \dots, m$, $j = 1, \dots, n$ elementos de F. Temos que existe uma única aplicação bilinear $B : E_1 \times E_2 \to F$ tal que:*

$$B(e_i^1, e_j^2) = f_{ij}, \ \text{para todos } i = 1, \dots, m, \ j = 1, \dots, n. \tag{A.1.2}$$

DEMONSTRAÇÃO. Para $i = 1, \dots, m$, seja $\pi_i^1 : E_1 \to \mathbb{K}$ o funcional linear que associa a cada $x \in E_1$ a sua i-ésima coordenada na base $\mathcal{B}^1$ e para cada $j = 1, \dots, n$, seja $\pi_j^2 : E_2 \to \mathbb{K}$ o funcional linear que associa a cada $y \in E_2$ a sua j-ésima coordenada na base $\mathcal{B}^2$. Vamos tentar descobrir como a aplicação bilinear desejada pode ser obtida: se $B : E_1 \times E_2 \to F$ é uma aplicação bilinear satisfazendo a condição (A.1.2) então:

$$B(x,y) = B\Big(\sum_{i=1}^{m} \pi_i^1(x)e_i^1, \sum_{j=1}^{n} \pi_j^2(y)e_j^2\Big) = \sum_{i=1}^{m}\sum_{j=1}^{n} \pi_i^1(x)\pi_j^2(y)B(e_i^1,e_j^2)$$
$$= \sum_{i=1}^{m}\sum_{j=1}^{n} \pi_i^1(x)\pi_j^2(y)f_{ij},$$

para todos $x \in E_1$, $y \in E_2$. A única opção que temos então para definir B é fazer:

$$B(x,y) = \sum_{i=1}^{m}\sum_{j=1}^{n} \pi_i^1(x)\pi_j^2(y)f_{ij}, \tag{A.1.3}$$

para todos $x \in E_1$, $y \in E_2$. A demonstração da proposição é obtida então através da verificação (puramente mecânica) de que a aplicação B definida por (A.1.3) é realmente bilinear e satisfaz a condição (A.1.2). □

A.1.12. Observação. Na Proposição A.1.11 estamos assumindo implicitamente que os espaços vetoriais E^1, E^2 têm dimensão finita, mas na verdade isso não é necessário: bastaria considerar bases indexadas em conjuntos de índices arbitrários (possivelmente infinitos). A demonstração do resultado em dimensão infinita seria exatamente igual, já que cada vetor tem no máximo um número finito de coordenadas não nulas numa dada base e todas as somatórias aparentemente infinitas que apareceriam na demonstração seriam na verdade somas de famílias *quase nulas*, isto é, famílias $(v_i)_{i\in I}$ em que $v_i \neq 0$ apenas para um número finito de índices $i \in I$.

A Proposição A.1.11 pode também ser lida assim: uma aplicação bilinear $B : E_1 \times E_2 \to F$ fica unicamente definida pelos seus valores $B(e_i^1, e_j^2)$, $i = 1,\ldots,m$, $j = 1,\ldots,n$, em vetores de bases $\mathcal{B}^1 = (e_i^1)_{i=1}^m$, $\mathcal{B}^2 = (e_i^2)_{i=1}^n$ de E_1 e E_2, respectivamente. Se escrevemos $B_{ij} = B(e_i^1, e_j^2) \in F$, $i = 1,\ldots,m$, $j = 1,\ldots,n$ então obtemos uma matriz $(B_{ij})_{m\times n}$ com entradas no espaço vetorial F que caracteriza completamente a aplicação bilinear B. No caso em que $F = \mathbb{K}$ então $(B_{ij})_{m\times n}$ é uma matriz de escalares que é normalmente conhecida como *a matriz da aplicação bilinear* B com respeito às bases $\mathcal{B}^1$, $\mathcal{B}^2$. No caso geral (F não necessariamente igual a $\mathbb{K}$) poderíamos escolher uma base $\mathcal{C} = (f_i)_{i=1}^p$ de F (vamos supor que F tem dimensão finita, por simplicidade) e denotar por B_{ij}^k a k-ésima coordenada de $B_{ij} \in F$ na base $\mathcal{C}$, para $i = 1,\ldots,m$, $j = 1,\ldots,n$, $k = 1,\ldots,p$. A aplicação bilinear $B : E_1 \times E_2 \to F$ fica então completamente caracterizada por uma matriz de três índices $(B_{ij}^k)_{m\times n\times p}$ com entradas em $\mathbb{K}$.

A.1.13. Exercício. Se E, E_1, E_2, F são espaços vetoriais de dimensão finita, determina as dimensões dos espaços $\mathrm{Lin}(E,F)$ e $\mathrm{Lin}(E_1,E_2;F)$ em função das dimensões de E, E_1, E_2 e F.

A noção de aplicação bilinear admite uma generalização natural, a saber:

A.1.14. Definição. Sejam $E_1, \ldots, E_n, F$ espaços vetoriais. Uma aplicação $B : E_1 \times \cdots \times E_n \to F$ é dita *n-linear* ou *multilinear* se for "linear em cada uma de suas variáveis"; mais precisamente, isso significa que para todo $i = 1, \ldots, n$ e para todos $x_1 \in E_1, \ldots, x_{i-1} \in E_{i-1}, x_{i+1} \in E_{i+1}, \ldots, x_n \in E_n$, a aplicação:

$$E_i \ni x_i \longmapsto B(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) \in F$$

é linear.

Evidentemente, a noção de aplicação n-linear reduz-se à noção de aplicação linear para $n = 1$ e reduz-se à noção de aplicação bilinear para $n = 2$. A noção de aplicação n-linear pode também ser descrita através de $2n$ igualdades (generalizando (A.1.1)).

A.1.15. Notação. $\mathrm{Lin}(E_1, \ldots, E_n; F)$ denota o conjunto das aplicações multilineares $B : E_1 \times \cdots \times E_n \to F$, onde $E_1, \ldots, E_n, F$ são espaços vetoriais. Se $E = E_1 = \cdots = E_n$, escrevemos também $\mathrm{Lin}_n(E; F)$.

A.1.16. Exercício. Dados espaços vetoriais $E_1, \ldots, E_n, F$, mostre que $\mathrm{Lin}(E_1, \ldots, E_n; F)$ é um espaço vetorial quando munido das operações definidas pelas igualdades:

$$\begin{gathered}(B + B')(x_1, \ldots, x_n) = B(x_1, \ldots, x_n) + B'(x_1, \ldots, x_n),\\ (\lambda B)(x_1, \ldots, x_n) = \lambda\big(B(x_1, \ldots, x_n)\big),\end{gathered}$$

onde $x_1 \in E_1, \ldots, x_n \in E_n$, $\lambda \in \mathbb{K}$, $B, B' \in \mathrm{Lin}(E_1, \ldots, E_n; F)$.

A.1.17. Exemplo. Dados números naturais $k_1, k_2, \ldots, k_{n+1}$ então a aplicação:

$$\begin{gathered}\mathrm{M}_{k_1\times k_2}(\mathbb{K}) \times \mathrm{M}_{k_2\times k_3}(\mathbb{K}) \times \cdots \times \mathrm{M}_{k_n\times k_{n+1}}(\mathbb{K}) \longrightarrow \mathrm{M}_{k_1\times k_{n+1}}(\mathbb{K})\\ (A_1, \ldots, A_n) \longmapsto A_1 \cdots A_n\end{gathered}$$

é n-linear.

A.1.18. Exemplo. A aplicação:

$$\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n \ni (x, y, z) \longmapsto \langle x, y\rangle z \in \mathbb{R}^n$$

é *trilinear* (isto é, 3-linear).

A.1.19. Exemplo. Dados espaços vetoriais E_1, E_2, F então a *aplicação de avaliação*:

$$\mathrm{aval} : \mathrm{Lin}(E_1, E_2; F) \times E_1 \times E_2 \ni (B, x_1, x_2) \longmapsto B(x_1, x_2) \in F$$

é trilinear. Mais geralmente, se são dados espaços vetoriais $E_1, \ldots, E_n$ então a aplicação:

$$\mathrm{Lin}(E_1, \ldots, E_n; F) \times E_1 \times \cdots \times E_n \ni (B, x_1, \ldots, x_n) \longmapsto B(x_1, \ldots, x_n) \in F$$

é $(n + 1)$-linear.

Um bom exemplo de aplicação multilinear é obtido a partir da noção de determinante de uma matriz. Recorde que se A é uma matriz $n \times n$ com entradas em $\mathbb{K}$ então o *determinante* de A é definido por:

$$\text{(A.1.4)} \qquad \det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)}.$$

Na fórmula acima S_n denota o conjunto (que é um grupo, mas isso não vem ao caso) de todas as *permutações de n elementos*, isto é, de todas as funções bijetoras $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. O *sinal* de σ, denotado por $\operatorname{sgn}(\sigma)$, é igual a 1 se σ é uma permutação par e igual a -1 se σ é uma permutação ímpar[3].

A.1.20. Exercício. Se $A \in \mathrm{M}_n(\mathbb{K})$, mostre que $\det(A) = \det(A^{\mathrm{t}})$, onde A^{t} denota a matriz transposta de A (sugestão: faça $\tau = \sigma^{-1}$ na somatória (A.1.4) e observe que $\prod_{i=1}^{n} A_{i\sigma(i)} = \prod_{i=1}^{n} A_{\tau(i)i}$ e que $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau)$).

A.1.21. Exemplo. Dados vetores $v_1, \ldots, v_n \in \mathbb{K}^n$, seja:

$$\det(v_1, \ldots, v_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) v_{1\sigma(1)} \cdots v_{n\sigma(n)}$$

o determinante da matriz $n \times n$ que possui os vetores $v_1, \ldots, v_n$ em suas colunas (ou, em vista do resultado do Exercício A.1.20, pode ser nas linhas também!). Temos então que $\det \in \mathrm{Lin}_n(\mathbb{K}^n; \mathbb{K})$.

As Proposições A.1.10 e A.1.11 (e suas demonstrações) podem ser generalizadas de forma mais ou menos óbvia para o contexto de aplicações multilineares. Colocamos aqui os enunciados e demonstrações resumidas; sua aparência é meio desajeitada, mas espero que isso não intimide o leitor.

A.1.22. Proposição. *Sejam $E_1, \ldots, E_n$, F espaços vetoriais, C_i um conjunto de geradores para E_i, $i = 1, \ldots, n$, e $B, B' \in \mathrm{Lin}(E_1, \ldots, E_n; F)$ aplicações n-lineares. Se:*

$$B(x_1, \ldots, x_n) = B'(x_1, \ldots, x_n),$$

para todos $x_1 \in C_1, \ldots, x_n \in C_n$ então $B = B'$.

Demonstração. Dados $x_i \in E_i$, $i = 1, \ldots, n$, escrevemos:

$$x_i = \sum_{\alpha=1}^{r_i} \lambda_i^{\alpha} x_i^{\alpha},$$

[3]Recorde que uma *transposição* é uma permutação que move exatamente dois elementos. Toda permutação pode ser decomposta num produto de transposições; tal decomposição não é única, mas a *paridade* do número de transposições utilizadas na decomposição não depende da decomposição escolhida. Dizemos então que uma permutação σ é par (resp., ímpar) se o número de transposições que aparece numa decomposição qualquer de σ é par (resp., ímpar).

com $x_i^\alpha \in C_i$, $\lambda_i^\alpha \in \mathbb{K}$, $\alpha = 1, \ldots, r_i$, $i = 1, \ldots, n$; daí:

$$\begin{aligned} B(x_1, \ldots, x_n) &= \sum_{\alpha_1=1}^{r_1} \cdots \sum_{\alpha_n=1}^{r_n} \lambda_1^{\alpha_1} \cdots \lambda_n^{\alpha_n} B(x_1^{\alpha_1}, \ldots, x_n^{\alpha_n}) \\ &= \sum_{\alpha_1=1}^{r_1} \cdots \sum_{\alpha_n=1}^{r_n} \lambda_1^{\alpha_1} \cdots \lambda_n^{\alpha_n} B'(x_1^{\alpha_1}, \ldots, x_n^{\alpha_n}) = B'(x_1, \ldots, x_n). \end{aligned}$$

Logo $B = B'$. □

A.1.23. Proposição. *Sejam E_1, ..., E_n, F espaços vetoriais, seja:*

$$\mathcal{B}^i = (e_\alpha^i)_{\alpha=1}^{m_i}$$

uma base de E_i, $i = 1, \ldots, n$ e sejam:

$$f_{\alpha_1 \ldots \alpha_n} \in F, \ \alpha_i = 1, \ldots, m_i, \ i = 1, \ldots, n$$

elementos de F. Temos que existe uma única aplicação n-linear

$$B : E_1 \times \cdots \times E_n \longrightarrow F$$

tal que:

(A.1.5) $B(e_{\alpha_1}^1, \ldots, e_{\alpha_n}^n) = f_{\alpha_1 \ldots \alpha_n}$, *para todos $\alpha_i = 1, \ldots, m_i$, $i = 1, \ldots, n$.*

Demonstração. Para $i = 1, \ldots, n$, $\alpha = 1, \ldots, m_i$, seja $\pi_\alpha^i : E_i \to \mathbb{K}$ o funcional linear que associa a cada elemento de E_i sua α-ésima coordenada na base $\mathcal{B}^i$. É fácil ver que a aplicação B definida por:

$$B(x_1, \ldots, x_n) = \sum_{\alpha_1=1}^{m_1} \cdots \sum_{\alpha_n=1}^{m_n} \pi_{\alpha_1}^1(x_1) \cdots \pi_{\alpha_n}^n(x_n) f_{\alpha_1 \ldots \alpha_n}, \tag{A.1.6}$$

para todos $x_1 \in E_1$, ..., $x_n \in E_n$ é n-linear e satisfaz a condição (A.1.5). A unicidade de B segue da Proposição A.1.22 (ou de um argumento para se chegar à fórmula (A.1.6) análogo ao que aparece na demonstração da Proposição A.1.11). □

A.1.24. Observação. A Proposição A.1.23 pode ser facilmente generalizada para o caso em que a dimensão dos espaços vetoriais E_1, ..., E_n não é necessariamente finita (veja Observação A.1.12).

A Proposição A.1.23 pode também ser lida assim: uma aplicação multilinear $B : E_1 \times \cdots \times E_n \to F$ fica unicamente definida pelos seus valores $B(e_{\alpha_1}^1, \ldots, e_{\alpha_n}^n)$, $\alpha_i = 1, \ldots, m_i$, $i = 1, \ldots, n$, em vetores de bases $\mathcal{B}^i = (e_\alpha^i)_{\alpha=1}^{m_i}$, $i = 1, \ldots, n$ dos espaços E_1, ..., E_n. Se escrevemos:

$$B_{\alpha_1 \ldots \alpha_n} = B(e_{\alpha_1}^1, \ldots, e_{\alpha_n}^n) \in F, \quad \alpha_i = 1, \ldots, m_i, \ i = 1, \ldots, n,$$

obtemos uma matriz (de n índices) $m_1 \times \cdots \times m_n$ com entradas em F que caracteriza completamente a aplicação multilinear B. Tomando uma base $\mathcal{C} = (f_i)_{i=1}^p$ de F (vamos supor que F tem dimensão finita, por simplicidade) e denotando por $B_{\alpha_1 \ldots \alpha_n}^j$ a j-ésima coordenada na base $\mathcal{C}$ de $B_{\alpha_1 \ldots \alpha_n}$, obtemos uma matriz (de $n+1$ índices) $m_1 \times \cdots \times m_n \times p$ com entradas em $\mathbb{K}$

que caracteriza completamente a aplicação multilinear B. Chamamos essa a *matriz que representa* B *com respeito à bases* $\mathcal{B}^1, \ldots, \mathcal{B}^n, \mathcal{C}$.

A.1.25. Exercício. Dados espaços vetoriais $E_1, \ldots, E_n$, F de dimensão finita, determine a dimensão de $\mathrm{Lin}(E_1, \ldots, E_n; F)$ em função das dimensões dos espaços E_i, $i = 1, \ldots, n$, e da dimensão de F.

A.1.1. Tensores. Nós às vezes chamaremos aplicações multilineares de *tensores*. A palavra "tensor" aparece em Matemática com muitos significados diferentes. Em alguns textos de Física, tensores são matrizes com vários índices que "transformam-se de modo adequado quando fazemos mudanças de base"; na verdade, essas matrizes de vários índices são as matrizes que representam aplicações multilineares com respeito à uma dada escolha de bases[4]. Tensores são importantes na Física para modelagem de objetos como campos eletromagnéticos e campos gravitacionais (na Teoria da Relatividade Geral). Em Geometria Diferencial (na Geometria Riemannian, por exemplo) são usados tensores para se representar objetos geométricos tais como a curvatura de uma variedade Riemanniana. A palavra "tensor" é muitas vezes usada também para se referir a um elemento de um *produto tensorial*; produto tensorial é uma construção algébrica[5] que nos dá um novo espaço vetorial $E_1 \otimes E_2$ (lê-se "E_1 tensor E_2") a partir de espaços vetorais E_1, E_2. Ocorre que no contexto de espaços vetoriais de dimensão finita há um isomorfismo natural entre produtos tensoriais e espaços de aplicações multilineares: a saber, $E_1 \otimes E_2$ é naturalmente isomorfo a $\mathrm{Lin}(E_1^*, E_2^*; \mathbb{K})$ e $\mathrm{Lin}(E_1, \ldots, E_n; F)$ é naturalmente isomorfo a $E_1^* \otimes \cdots \otimes E_n^* \otimes F$. Notamos que tais isomorfismos não existem no contexto geral de espaços vetorias de dimensão possivelmente infinita (ou no contexto geral de módulos), mas apesar disso nós usaremos a palavra "tensor" como sinônimo de "aplicação multilinear".

A.2. Soma direta

Se E é um espaço vetorial e se E_1, E_2 são subespaços de E denotamos por $E_1 + E_2$ o conjunto de todas as somas $x_1 + x_2$, com $x_1 \in E_1$, $x_2 \in E_2$;

[4]Particularmente importantes são os chamados (p, q)-*tensores* ou *tensores* p *vezes covariantes e* q *vezes contravariantes* num dado espaço vetorial E; tratam-se de aplicações multilineares de $E \times \cdots \times E \times E^* \times \cdots \times E^*$ (p fatores E e q fatores E^*) em $\mathbb{K}$. Escolhendo uma base em E e tomando a base dual em E^*, essas aplicações multilineares são representadas por matrizes de $p + q$ índices; normalmente escreve-se os p índices correspondentes às cópias de E embaixo (os *índices covariantes*) e os q índices correspondentes às cópias de E^* em cima (os *índices contravariantes*).

[5]Um *produto tensorial* de espaços vetoriais E_1, E_2 é um espaço vetorial $E_1 \otimes E_2$ junto com uma aplicação bilinear $E_1 \times E_2 \ni (x, y) \mapsto x \otimes y \in E_1 \otimes E_2$ tal que para todo espaço vetorial F e toda aplicação bilinear $B : E_1 \times E_2 \to F$ existe uma única aplicação linear $\widetilde{B} : E_1 \otimes E_2 \to F$ tal que $\widetilde{B}(x \otimes y) = B(x, y)$, para todos $x \in E_1$, $y \in E_2$. Mostra-se que, a menos de isomorfismos (num sentido adequado), dois espaços vetoriais admitem um único produto tensorial. Essa construção pode também ser generalizada para o contexto de módulos sobre anéis arbitrários, embora algumas adaptações sejam necessárias quando o anel de escalares não é comutativo.

claramente, $E_1 + E_2$ coincide com o subespaço de E gerado por $E_1 \cup E_2$. Se $E_1 \cap E_2 = \{0\}$, dizemos que a soma de E_1 com E_2 é *direta* e denotamos o subespaço $E_1 + E_2$ por $E_1 \oplus E_2$. Evidentemente, $E = E_1 + E_2$ se e somente se todo vetor de E é soma de um vetor de E_1 com um vetor de E_2; além do mais, é fácil ver que $E = E_1 \oplus E_2$ se e somente se todo vetor de E se escreve *de modo único* como soma de um vetor de E_1 com um vetor de E_2. Generalizando essas idéias para o contexto de um número finito arbitrário[6] de subespaços de E, obtemos a seguinte:

A.2.1. Proposição. *Seja E um espaço vetorial e $E_1, \ldots, E_n$ subespaços de E. São equivalentes:*

(a) *para todo $i = 1, \ldots, n$, a interseção:*

$$E_i \cap \sum_{\substack{j=1 \\ j \neq i}}^{n} E_j \tag{A.2.1}$$

contém somente o vetor nulo;

(b) *para todos $x_1 \in E_1, \ldots, x_n \in E_n$, se $x_1 + \cdots + x_n = 0$ então $x_1 = \cdots = x_n = 0$;*

(c) *para todos $x_1, y_1 \in E_1, \ldots, x_n, y_n \in E_n$, se:*

$$x_1 + \cdots + x_n = y_1 + \cdots + y_n$$

então $x_1 = y_1, \ldots, x_n = y_n$.

Demonstração. Deixamos a prova da equivalência entre (b) e (c) a cargo do leitor. Vamos provar a equivalência entre (a) e (b). Suponha (a). Se $x_1 + \cdots + x_n = 0$ e $x_i \in E_i$, $i = 1, \ldots, n$, então:

$$x_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} (-x_j) \in E_i \cap \sum_{\substack{j=1 \\ j \neq i}}^{n} E_j = \{0\},$$

donde $x_i = 0$, para todo $i = 1, \ldots, n$. Agora suponha (b). Dado $i = 1, \ldots, n$ e x na interseção (A.2.1) então existem $x_j \in E_j$, $j = 1, \ldots, n$, $j \neq i$, tais que:

$$x = \sum_{\substack{j=1 \\ j \neq i}}^{n} x_j.$$

Tomando $x_i = -x$ então $x_i \in E_i$ e $\sum_{j=1}^{n} x_j = 0$, donde $x_j = 0$, para todo $j = 1, \ldots, n$, sendo em particular $x = -x_i = 0$. □

[6] Dados subespaços $E_1, \ldots, E_n$ de E, então a soma $\sum_{i=1}^{n} E_i = E_1 + \cdots + E_n$ pode ser definida ou iterando a operação binária de soma de subespaços, ou tomando o conjunto de todas as somas $x_1 + \cdots + x_n$, com $x_1 \in E_1, \ldots, x_n \in E_n$, ou tomando o subespaço gerado pela união $\bigcup_{i=1}^{n} E_i$.

Quando uma das condições (a), (b), (c) que aparece no enunciado da Proposição A.2.1 ocorre dizemos que a soma dos subespaços $E_1, \ldots, E_n$ é *direta* e escrevemos $\bigoplus_{i=1}^n E_i$ ou $E_1 \oplus \cdots \oplus E_n$ para denotar a soma $\sum_{i=1}^n E_i$ (mas é bom observar que a notação $E_1 \oplus \cdots \oplus E_n$ é um pouco enganosa, já que a condição de a soma dos subespaços $E_1, \ldots, E_n$ ser direta é uma condição que envolve os n subespaços simultaneamente e não é equivalente à condição das somas $E_i + E_{i+1}$, $i = 1, \ldots, n-1$ serem diretas — e nem mesmo equivalente à condição mais forte de as somas $E_i + E_j$, $i, j = 1, \ldots, n$, $i \neq j$, serem diretas).

Note que a condição $E = \bigoplus_{i=1}^n E_i$ é equivalente à condição de que cada elemento $x \in E$ se escreva de modo único como uma soma:

$$x_1 + \cdots + x_n,$$

com $x_1 \in E_1, \ldots, x_n \in E_n$; chamamos a x_i a *componente de x no subespaço E_i* relativa à decomposição em soma direta $E = \bigoplus_{i=1}^n E_i$ (note que a componente de x no subespaço E_i depende não apenas de E_i mas dos outros E_j; por exemplo, se $E = E_1 \oplus E_2 = E_1 \oplus E_2'$ então é bem possível que a componente em E_1 de um elemento $x \in E$ com respeito à decomposição $E = E_1 \oplus E_2$ seja diferente da componente em E_1 de x com respeito à decomposição $E = E_1 \oplus E_2'$). A aplicação $\pi_i : E \to E_i$ que associa a cada $x \in E$ a sua componente em E_i relativa à decomposição $E = \bigoplus_{i=1}^n E_i$ é chamada a *projeção em E_i* relativa à decomposição $E = \bigoplus_{i=1}^n E_i$.

A.2.2. Exercício. Suponha que $E = \bigoplus_{i=1}^n E_i$. Mostre que:

(a) a aplicação $\pi_i : E \to E_i$ de projeção em E_i relativa à decomposição $E = \bigoplus_{i=1}^n E_i$ é linear;

(b) mostre que para todo $i = 1, \ldots, n$ a restrição de π_i a E_i é igual à aplicação identidade de E_i e que para $j = 1, \ldots, n$, $j \neq i$, a restrição de π_i a E_j é a aplicação nula;

(c) trocando o contra-domínio da aplicação π_i de E_i para E (de modo que π_i seja um elemento de $\mathrm{Lin}(E, E)$), mostre que π_i^2 (isto é, $\pi_i \circ \pi_i$) é igual a π_i, que $\pi_i \circ \pi_j = 0$ para $i, j = 1, \ldots, n$, $i \neq j$, e que $\sum_{i=1}^n \pi_i$ é igual à aplicação identidade de E.

A.2.3. Exemplo. Se $(e_i)_{i=1}^n$ denota a base canônica de $\mathbb{K}^n$ e se $\mathbb{K}e_i$ denota o subespaço gerado pelo vetor e_i então $\mathbb{K}^n = \bigoplus_{i=1}^n \mathbb{K}e_i$.

A.2.4. Exercício. Dado um espaço vetorial E, uma base[7] $\mathcal{B}$ de E e uma partição $\mathcal{B} = \bigcup_{i=1}^n \mathcal{B}_i$ de $\mathcal{B}$ (isto é, $\mathcal{B}_i \cap \mathcal{B}_j = \emptyset$, para $i = 1, \ldots, n$, $i \neq j$) então, denotando por E_i o subespaço gerado por $\mathcal{B}_i$, mostre que $E = \bigoplus_{i=1}^n E_i$.

[7]Para mim uma base é uma *família* (indexada) de elementos e não apenas um conjunto; a distinção entre família e conjunto é importante: por exemplo, no caso de uma base de um espaço vetorial de dimensão finita precisamos de uma *ordem* na base, para que se possa falar em primeira, segunda, etc., coordenada de um vetor. No entanto, às vezes cometo um certo abuso e trato uma base como se fosse apenas um conjunto, para facilitar a notação.

A.2.5. Exercício. Se $E = \bigoplus_{i=1}^n E_i$ e $\mathcal{B}_i$ é uma base de E_i, $i = 1, \ldots, n$, mostre que $\mathcal{B}_i \cap \mathcal{B}_j = \emptyset$, para $i, j = 1, \ldots, n$, $i \neq j$ e que $\bigcup_{i=1}^n \mathcal{B}_i$ é uma base de E.

As Proposições A.2.6 e A.2.7 abaixo são ferramentas práticas para se definir aplicações lineares com domínio ou contra-domínio numa soma direta[8].

A.2.6. Proposição. *Se $E = \bigoplus_{i=1}^n E_i$, F são espaços vetoriais e se $T_i : E_i \to F$, $i = 1, \ldots, n$, são aplicações lineares então existe uma única aplicação linear $T : E \to F$ tal que $T|_{E_i} = T_i$, para $i = 1, \ldots, n$.*

Demonstração. Se $T : E \to F$, $T' : E \to F$ são aplicações lineares tais que $T|_{E_i} = T'|_{E_i} = T_i$, para $i = 1, \ldots, n$, então T e T' concidem em $\bigcup_{i=1}^n E_i$, que é um conjunto de geradores de E; logo $T = T'$. A existência de T é demonstrada tomando $T = \sum_{i=1}^n T_i \circ \pi_i$, onde $\pi_i : E \to E_i$ denota a projeção em E_i relativa à decomposição $E = \bigoplus_{i=1}^n E_i$. □

A.2.7. Proposição. *Se $E = \bigoplus_{i=1}^n E_i$, F são espaços vetoriais e se $T_i : F \to E_i$, $i = 1, \ldots, n$, são aplicações lineares então existe uma única aplicação linear $T : F \to E$ tal que $\pi_i \circ T = T_i$, para $i = 1, \ldots, n$, onde $\pi_i : E \to E_i$ denota a projeção em E_i relativa à decomposição $E = \bigoplus_{i=1}^n E_i$.*

Demonstração. Se $\theta_i : E_i \to E$ denota a aplicação inclusão então (Exercício A.2.2, item (c)) $\sum_{i=1}^n \theta_i \circ \pi_i$ é a aplicação identidade de E, de modo que as igualdades $\pi_i \circ T = T_i$, $i = 1, \ldots, n$, implicam em:

$$T = \sum_{i=1}^n \theta_i \circ T_i. \tag{A.2.2}$$

Isso mostra que a aplicação T é única, se existir (obrigatoriamente dada pela fórmula (A.2.2)). Definindo T usando (A.2.2), obtemos:

$$\pi_j \circ T = \sum_{i=1}^n \pi_j \circ \theta_i \circ T_i = T_j,$$

já que $\pi_j \circ \theta_j$ é igual à aplicação identidade de E_j e $\pi_j \circ \theta_i = 0$ para $i \neq j$. □

A.2.8. Observação. É possível também definir somas diretas para famílias possivelmente infinitas $(E_i)_{i \in I}$ de subespaços de um espaço vetorial E. Quando se trabalha com famílias infinitas, somas do tipo $\sum_{i \in I} x_i$, $x_i \in E_i$, fazem sentido apenas quando $(x_i)_{i \in I}$ é uma *família quase nula* (veja Observação A.1.12). Todos os resultados apresentados nesta seção generalizam-se facilmente para o caso de somas diretas infinitas, exceto a Proposição A.2.7, que é falsa para somas diretas infinitas.

---

[8]Essas proposições exprimem fatos que têm um significado especial quando se estuda Teoria das Categorias: a saber, a Proposição A.2.6 exprime o fato que somas diretas $\bigoplus_{i=1}^n E_i$ são *somas* na categoria de espaços vetoriais e aplicações lineares e a Proposição A.2.7 exprime o fato que somas diretas $\bigoplus_{i=1}^n E_i$ são *produtos* nessa categoria.

A.2.1. Somas diretas internas e externas. Dados espaços vetoriais $E_1, \ldots, E_n$ (não necessariamente subespaços de um mesmo espaço vetorial) então podemos tornar o produto cartesiano $E = \prod_{i=1}^n E_i = E_1 \times \cdots \times E_n$ um espaço vetorial definindo operações:

$$\begin{gathered}(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n),\\ \lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n),\end{gathered} \tag{A.2.3}$$

para $x_i, y_i \in E_i$, $i = 1, \ldots, n$, $\lambda \in \mathbb{K}$. A aplicação $\theta_i : E_i \to E$ definida por:

$$\theta_i(x) = (0, \ldots, 0, x, 0, \ldots, 0) \quad \text{(com o } x \text{ na } i\text{-ésima coordenada)},$$

para todo $x \in E_i$, é linear injetora e nos dá um isomorfismo de E_i sobre o subespaço $\theta_i(E_i) = \{0\}^{i-1} \times E_i \times \{0\}^{n-i}$ de E. É fácil ver que:

$$E = \bigoplus_{i=1}^n \theta_i(E_i)$$

e se identificarmos E_i com $\theta_i(E_i)$ escreveremos apenas $E = \bigoplus_{i=1}^n E_i$. É comum chamar ao espaço vetorial $\prod_{i=1}^n E_i$ (munido das operações definidas em (A.2.3)) a *soma direta externa* dos espaços $E_1, \ldots, E_n$; daí, a soma direta discutida no início da Seção A.2 (em que os E_i devem ser todos subespaços de um mesmo espaço E) é também chamada *soma direta interna* dos espaços vetoriais $E_1, \ldots, E_n$. Observe que se $E_1, \ldots, E_n$ são subespaços de um mesmo espaço vetorial E então também faz sentido considerar a soma direta externa $\prod_{i=1}^n E_i$, não importando se a soma $\sum_{i=1}^n E_i$ é direta ou não! No entanto, se a soma $\sum_{i=1}^n E_i$ for direta, então é fácil ver que a aplicação:

$$\prod_{i=1}^n E_i \ni (x_1, \ldots, x_n) \longmapsto \sum_{i=1}^n x_i \in \bigoplus_{i=1}^n E_i \tag{A.2.4}$$

é um isomorfismo linear. Vimos então que:

(a) a soma direta externa $\prod_{i=1}^n E_i$ é igual à soma direta interna de espaços $\theta_i(E_i)$ isomorfos aos E_i (o isomorfismo entre E_i e $\theta_i(E_i)$ corresponde apenas ao acréscimo de algumas coordenadas zero);

(b) se $E = \bigoplus_{i=1}^n E_i$ é soma direta interna de subespaços E_i então (A.2.4) nos dá um isomorfismo entre E e a soma direta externa $\prod_{i=1}^n E_i$.

Em vista de (a) e (b) é comum simplesmente ignorar a diferença entre soma direta interna e externa. Nós adotaremos essa prática, sempre que possível.

A.2.9. Exercício. Se um espaço vetorial E é soma direta (interna) de subespaços $E_1, \ldots, E_n$, $\pi_i : E \to E_i$ denota a projeção em E_i relativa à decomposição $E = \bigoplus_{i=1}^n E_i$ e $\bar{\pi}_i : \prod_{i=1}^n E_i \to E_i$ denota a i-ésima projeção do produto cartesiano $\prod_{i=1}^n E_i$, mostre que o diagrama:

$$\begin{tikzcd} \prod_{i=1}^n E_i \arrow[rr, "(A.2.4)", "\cong"'] \arrow[dr, "\bar{\pi}_i"'] & & E \arrow[dl, "\pi_i"] \\ & E_i & \end{tikzcd}$$

é comutativo, para todo $i = 1, \ldots, n$, isto é, se θ denota o isomorfismo (A.2.4) então $\pi_i \circ \theta = \bar{\pi}_i$, para todo $i = 1, \ldots, n$.

A.2.10. Observação. Se $(E_i)_{i \in I}$ é uma família infinita de espaços vetorais então o produto cartesiano $\prod_{i \in I} E_i$ em geral não é soma direta interna dos subespaços $\theta_i(E_i)$. Por isso, a *soma direta externa* é, nesse contexto, definida como sendo o subespaço de $\prod_{i \in I} E_i$ formado pelas famílias quase nulas; esse subespaço é igual à soma direta interna dos subespaços $\theta_i(E_i)$. Evidentemente, no caso em que I é finito qualquer família $(x_i)_{i \in I}$ é quase nula. Se I é infinito e $E = \bigoplus_{i \in I} E_i$ (soma direta interna) então a aplicação $\prod_{i \in I} E_i \ni (x_i)_{i \in I} \mapsto \sum_{i \in I} x_i \in E$ não está em geral sequer bem definida, mas se trocamos $\prod_{i \in I} E_i$ pelo espaço de famílias quase nulas então essa aplicação torna-se um isomorfismo linear. É interessante notar que, embora a Proposição A.2.7 não seja verdadeira para famílias infinitas de espaços vetoriais (veja Observação A.2.8), ela tornaria-se verdaderia[9] se trocássemos a hipótese $E = \bigoplus_{i \in I} E_i$ pela hipótese $E = \prod_{i \in I} E_i$.

Exercícios para o Apêndice A

A.1. Exercício. Seja E um espaço vetorial real. Um *produto interno* em E é uma aplicação bilinear:

$$E \times E \ni (x, y) \longmapsto \langle x, y \rangle \in \mathbb{R}$$

tal que $\langle x, y \rangle = \langle y, x \rangle$, para todos $x, y \in E$ e $\langle x, x \rangle > 0$, para todo $x \in E$ não nulo. Se $\langle \cdot, \cdot \rangle$ é um produto interno em E, nós definimos:

$$\|x\| = \langle x, x \rangle^{\frac{1}{2}}, \quad d(x, y) = \|x - y\|,$$

para todos $x, y \in E$. Mostre que:

(a) para todos $x, y \in E$, vale a *desigualdade de Cauchy–Schwarz*:

$$|\langle x, y \rangle| \le \|x\| \, \|y\|,$$

sendo que a igualdade vale se e somente se x, y são linearmente dependentes (sugestão: $p(t) = \langle x + ty, x + ty \rangle \ge 0$, para todo $t \in \mathbb{R}$; analise o sinal do discriminante do polinômio do segundo grau p);

[9] Para quem conhece um pouco de teoria das categorias, o que está ocorrendo aqui pode ser entendido da seguinte maneira: para famílias finitas de objetos, somas e produtos coincidem na categoria de espaços vetoriais e aplicações lineares. No entanto, o mesmo não vale para famílias infinitas de objetos.

(b) para todos $x, y \in E$, vale a *desigualdade triangular*:

$$\|x + y\| \leq \|x\| + \|y\|. \tag{A.5}$$

(sugestão: calcule $\langle x + y, x + y\rangle$ e use a desigualdade de Cauchy–Schwarz);

(c) para todos $x, y, z \in E$ vale também a seguinte *desigualdade triangular*:

$$d(x, z) \leq d(x, y) + d(y, z).$$

(sugestão: $x - z = (x - y) + (y - z)$; use (A.5)).

APÊNDICE B

Espaços métricos

Este apêndice contém um curso relâmpago sobre a teoria básica de espaços métricos. A exposição é feita de forma bem sucinta, de modo que o apêndice provavelmente não será adequado para leitores que estão tomando contato com a teoria de espaços métricos pela primeira vez. Todas as definições e teoremas (com demonstração!) que considero relevantes como ferramentas para um curso de Cálculo no $\mathbb{R}^n$ são apresentadas. Alguns resultados que considero simples de demonstrar são deixados a cargo do leitor sob a forma de exercícios ao longo do texto; para o leitor menos experiente, resolver esses exercícios é fundamental para que o mesmo adquira alguma intimidade com o assunto. Observo que alguns desses exercícios são quase triviais e para resolvê-los basta compreender as definições dos conceitos envolvidos.

B.1. Definição e conceitos básicos

B.1.1. Definição. Seja M um conjunto. Uma *métrica* em M é uma função $d : M \times M \to \mathbb{R}$ que satisfaz as seguintes condições:

(a) $d(x, y) \geq 0$, para todos $x, y \in M$;
(b) para todos $x, y \in M$, $d(x, y) = 0$ se e somente se $x = y$;
(b) $d(x, y) = d(y, x)$, para todos $x, y \in M$;
(c) (*desigualdade triangular*) $d(x, z) \leq d(x, y) + d(y, z)$, para todos $x, y, z \in M$.

Se d é uma métrica em M então o par (M, d) é chamado um *espaço métrico*.

Para deixar a linguagem menos carregada, nós muitas vezes escrevemos frases como "seja M um espaço métrico..." em vez de "seja (M, d) um espaço métrico..."; às vezes dizemos também que M está *munido* de uma métrica d, significando que referências ao "espaço métrico M" devem ser entendidas como referências ao "espaço métrico (M, d)".

B.1.2. Exemplo. A *métrica Euclideana* em $\mathbb{R}^n$ é definida por:

$$d(x, y) = \Big(\sum_{i=1}^{n} (x_i - y_i)^2 \Big)^{\frac{1}{2}},$$

para todos $x, y \in \mathbb{R}^n$. Deixamos a cargo do leitor a verificação do fato que d é realmente uma métrica (a desigualdade triangular para d é mais difícil, mas segue diretamente do resultado do Exercício A.1 aplicado ao produto interno canônico de $\mathbb{R}^n$). A menos de menção explícita em contrário, consideramos

$\mathbb{R}^n$ sempre munido da métrica Euclideana. Também são métricas em $\mathbb{R}^n$ as funções:

$$d_1 : \mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \longmapsto \sum_{i=1}^{n} |x_i - y_i| \in \mathbb{R},$$
$$d_\infty : \mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \longmapsto \max\big\{|x_i - y_i| : i = 1, \ldots, n\big\} \in \mathbb{R}.$$

A métrica d_1 é também conhecida como *métrica da soma* e a métrica d_∞ é também conhecida como *métrica do máximo*. A métrica Euclideana d é às vezes também denotada por d_2. No caso particular em que $n = 1$, todas essas métricas coincidem:

$$d(x, y) = d_1(x, y) = d_\infty(x, y) = |x - y|, \quad x, y \in \mathbb{R}.$$

Veremos adiante (Seção B.5) que as métricas d, d_1, d_∞ são *equivalentes* num sentido que será esclarecido naquela seção.

B.1.3. Exemplo. Dado um conjunto arbitrário M, temos que a função $d : M \times M \to \mathbb{R}$ definida por:

$$d(x, y) = \begin{cases} 0, & \text{se } x = y, \\ 1, & \text{se } x \neq y, \end{cases}$$

para todos $x, y \in M$, é uma métrica em M chamada a *métrica zero-um*.

B.1.4. Exemplo. Se d é uma métrica em M e N é um subconjunto de M então a restrição $d|_{N\times N}$ é uma métrica em N chamada a *métrica induzida* por d em N. Dizemos também que $(N, d|_{N\times N})$ é um *subespaço* do espaço métrico (M, d). A menos de menção explícita em contrário, assume-se que um subconjunto de um espaço métrico está sempre munido da métrica induzida.

B.1.5. Definição. Dados um espaço métrico M, um ponto $x \in M$ e um número real $r > 0$ então a *bola aberta* de centro x e raio r é definida por:

$$\mathrm{B}(x, r) = \big\{y \in M : d(y, x) < r\big\}$$

e a *bola fechada* de centro x e raio r é definida por:

$$\mathrm{B}[x, r] = \big\{y \in M : d(y, x) \leq r\big\}.$$

Embora as notações $\mathrm{B}(x, r)$ e $\mathrm{B}[x, r]$ só façam referência explícita ao centro e ao raio das bolas, é evidente que as bolas $\mathrm{B}(x, r)$ e $\mathrm{B}[x, r]$ dependem também do conjunto M e da métrica d (que normalmente estão subentendidos pelo contexto). Em algumas situações nós temos que lidar com as bolas abertas e fechadas de vários espaços métricos e aí podem ser necessários ajustes de notação e terminologia para remover ambigüidades na comunicação. Esse é o caso do enunciado do Exercício B.1.6 a seguir, onde pedimos ao leitor para destrinchar a (fácil!) relação entre as bolas abertas e fechadas de um espaço métrico e as bolas abertas e fechadas de um subespaço desse espaço métrico.

B.1.6. Exercício. Sejam (M, d) um espaço métrico, N um subconjunto de M e considere o espaço métrico $(N, d|_{N\times N})$. Dados $x \in M$, $r > 0$, vamos denotar por $\mathrm{B}_M(x, r)$, $\mathrm{B}_M[x, r]$, respectivamente, a bola aberta e a bola fechada de centro x e raio r no espaço métrico (M, d); se $x \in N$, vamos denotar por $\mathrm{B}_N(x, r)$, $\mathrm{B}_N[x, r]$, respectivamente, a bola aberta e a bola fechada de centro x e raio r no espaço métrico $(N, d|_{N\times N})$. Mostre que, para todos $x \in N$, $r > 0$:

$$\mathrm{B}_N(x, r) = \mathrm{B}_M(x, r) \cap N, \quad \mathrm{B}_N[x, r] = \mathrm{B}_M[x, r] \cap N.$$

B.1.7. Definição. Dados um espaço métrico M e um subconjunto A de M então um ponto x de M é dito:

- *interior* a A se existe $r > 0$ tal que $\mathrm{B}(x, r) \subset A$;
- *exterior* a A se existe $r > 0$ tal que $\mathrm{B}(x, r) \cap A = \emptyset$, isto é, tal que $\mathrm{B}(x, r)$ está contido em $M \setminus A$ (o *complementar* de A em M);
- *um ponto de fronteira* de A se para todo $r > 0$ temos $\mathrm{B}(x, r) \cap A \neq \emptyset$ e $\mathrm{B}(x, r) \cap (M \setminus A) \neq \emptyset$;
- *aderente* a A se para todo $r > 0$ temos que $\mathrm{B}(x, r) \cap A \neq \emptyset$.

O *interior* de A, denotado por $\mathrm{int}(A)$ ou $\overset{\circ}{A}$, é definido como sendo o conjunto dos pontos interiores a A e a *fronteira* de A, denotada por ∂A, é definida como sendo o conjunto dos pontos de fronteira de A. O *fecho* de A, denotado por $\overline{A}$, é definido como sendo o conjunto dos pontos aderentes a A.

B.1.8. Exercício. Se A é um subconjunto de um espaço métrico M, mostre que o conjunto dos pontos exteriores a A coincide com o interior do complementar de A, isto é, com $\mathrm{int}(M \setminus A)$.

Em vista do resultado do Exercício B.1.8, nós não introduziremos notação (ou nome) para o conjunto dos pontos exteriores a A: nós simplesmente usaremos $\mathrm{int}(M \setminus A)$.

B.1.9. Exercício. Se A é um subconjunto de um espaço métrico M, mostre que todo ponto interior a A pertence a A (isto é, $\mathrm{int}(A)$ está contido em A), que todo ponto exterior a A pertence a $M \setminus A$ e que todo ponto de A é aderente a A (isto é, A está contido em $\overline{A}$). Tomando $M = \mathbb{R}$ munido da métrica Euclideana, dê exemplos de conjuntos A tais que:

(i) $\partial A \subset A$;
(ii) $\partial A \cap A = \emptyset$;
(iii) nem (i) nem (ii) ocorrem.

B.1.10. Exercício. Se A é um subconjunto de um espaço métrico M, mostre que as possibilidades "x é interior a A", "x é exterior a A" e "x é um ponto de fronteira de A" são exaustivas e mutuamente exclusivas para um dado ponto $x \in M$, isto é, mostre que os conjuntos $\mathrm{int}(A)$, ∂A e $\mathrm{int}(M \setminus A)$ cobrem M e são dois a dois disjuntos.

B.1.11. Exercício. Verifique que na Definição B.1.7 teria dado na mesma se tivéssemos usado bolas fechadas em vez de bolas abertas, isto é, obteríamos definições equivalentes se trocássemos todas as ocorrências de

$\mathrm{B}(x,r)$ por $\mathrm{B}[x,r]$ (sugestão: a bola fechada $\mathrm{B}\left[x,\frac{r}{2}\right]$ está contida na bola aberta $\mathrm{B}(x,r)$).

B.1.12. Exercício. Se A é um subconjunto de um espaço métrico M, mostre que:

(a) A e $M \setminus A$ têm a mesma fronteira, isto é, $\partial A = \partial(M \setminus A)$;
(b) para todo $x \in M$, x é um ponto de fronteira de A se e somente se x é simultaneamente aderente a A e a seu complementar $M \setminus A$. Em outras palavras:

$$\partial A = \overline{A} \cap \overline{M \setminus A}.$$

B.1.13. Exercício. Se A é um subconjunto de um espaço métrico M, mostre que, dado $x \in M$, são equivalentes as condições:

(i) x é aderente a A;
(ii) x não é um ponto exterior a A;
(iii) x é interior a A ou x é um ponto de fronteira de A;
(iv) x pertence a A ou x é um ponto de fronteira de A.

Conclua que:

$$\overline{A} = M \setminus \operatorname{int}(M \setminus A) = \operatorname{int}(A) \cup \partial A = A \cup \partial A. \tag{B.1.1}$$

B.1.14. Exercício. Se A é um subconjunto de um espaço métrico M, mostre que um dado ponto $x \in M$ é interior a A se e somente se não está no fecho de $M \setminus A$. Em outras palavras:

$$\operatorname{int}(A) = M \setminus \overline{M \setminus A}.$$

B.1.15. Definição. Um subconjunto A de um espaço métrico M é dito *denso* (ou *denso em M*) se todo ponto de M é aderente a A, isto é, se $\overline{A} = M$.

B.1.16. Exercício. Mostre que A é denso num espaço métrico M se e somente se $M \setminus A$ tem interior vazio.

B.1.17. Exercício. Dados subconjuntos A, B de um espaço métrico M, mostre que:

(a) se $A \subset B$ então $\operatorname{int}(A) \subset \operatorname{int}(B)$;
(b) se $A \subset B$ então $\overline{A} \subset \overline{B}$.

Evidentemente as noções de ponto interior, ponto exterior, ponto de fronteira, ponto aderente, interior, fecho e conjunto denso dependem do espaço métrico (M,d) no qual se está trabalhando: se (M,d) é um espaço métrico, N é um subconjunto de M, A é um subconjunto de N e $x \in A$ então é possível que x seja um ponto interior de A relativamente ao espaço métrico $(N, d|_{N\times N})$, mas que x não seja um ponto interior de A relativamente ao espaço métrico (M,d). Por exemplo, se $A = N$ então *todo* ponto $x \in A$ é interior a A relativamente ao espaço métrico $(N, d|_{N\times N})$ (pois uma bola aberta do espaço métrico $(N, d|_{N\times N})$ tem que estar contida em $A = N$!), mas evidentemente é possível que existam pontos $x \in A$ que não sejam interiores a A relativamente ao espaço métrico (M,d). Como usualmente o espaço métrico (M, d) relativamente ao qual os conceitos são considerados está subentendido pelo contexto, podemos escrever expressões como $\text{int}(A)$, $\overline{A}$ ou ∂A (ou falar em pontos interiores, aderentes ou de fronteira de A) sem fazer referência explícita ao espaço métrico (M, d); no entanto, quando o contexto não deixa claro o espaço métrico em questão pode ser necessário adotar ajustes de notação e terminologia que nos permitam explicitá-lo (é o que fizemos no enunciado do Exercício B.1.6). No Exercício B.1.18 a seguir pedimos ao leitor para olhar mais de perto a relação entre o fecho de um conjunto A num espaço métrico (M, d) e o fecho de A num subespaço de (M, d) que contém A.

B.1.18. Exercício. Sejam (M, d) um espaço métrico, N um subconjunto de M e considere o espaço métrico $(N, d|_{N\times N})$. Dados um subconjunto A de N e um ponto $x \in N$, mostre que x é aderente a A relativamente ao espaço métrico $(N, d|_{N\times N})$ se e somente se x é aderente a A relativamente ao espaço métrico (M, d) (sugestão: use o resultado do Exercício B.1.6). Conclua que o fecho de A relativamente a N é igual à interseção com N do fecho de A relativamente a M.

B.1.19. Definição. Seja M um espaço métrico. Um subconjunto A de M é dito *aberto* (ou *aberto em* M) se coincide com seu interior, isto é, se $\text{int}(A) = A$. Dizemos que A é *fechado* (ou *fechado em* M) se o seu complementar $M \setminus A$ é um conjunto aberto.

Como é sempre o caso que $\text{int}(A) \subset A$, temos que A é aberto se e somente se todo ponto de A é interior a A, ou ainda, se para todo $x \in A$ existe $r > 0$ tal que $\text{B}(x, r) \subset A$.

Note que “fechado” *não* é sinônimo de “não-aberto”!

B.1.20. Exercício. Mostre que um subconjunto A de um espaço métrico M é aberto se e somente se não contém nenhum dos seus pontos de fronteira, isto é, se e somente se $(\partial A) \cap A = \emptyset$.

B.1.21. Exercício. Seja A um subconjunto de um espaço métrico M. Mostre que:

(a) A é fechado se e somente se todo ponto de $M \setminus A$ é exterior a A (isto é, A é fechado se e somente se para todo $x \in M \setminus A$, existe $r > 0$ tal que $\text{B}(x, r) \cap A = \emptyset$);
(b) A é fechado se e somente se contém todos os seus pontos aderentes (note que, como é sempre o caso que $A \subset \overline{A}$, pode-se concluir daí que A é fechado se e somente se $A = \overline{A}$, isto é, *um conjunto é fechado se e somente se coincide com seu fecho*);
(c) A é fechado se e somente se contém todos os seus pontos de fronteira (isto é, A é fechado se e somente se $\partial A \subset A$).

B.1.22. Exercício. Mostre que todo subconjunto unitário de um espaço métrico M é um conjunto fechado (sugestão: dados $x \in M$ e $y \in M$, $y \neq x$, considere a bola aberta de centro y e raio $d(y, x) > 0$).

B.1.23. Exercício. Se M é um espaço métrico, U é um subconjunto aberto de M e A é um subconjunto arbitrário de M, mostre que U intercepta A se e somente se U intercepta $\overline{A}$.

B.1.24. Exercício. Mostre que um subconjunto A de um espaço métrico M é ao mesmo tempo aberto e fechado se e somente se sua fronteira é vazia.

B.1.25. Exercício. Dado um subconjunto A de um espaço métrico M, mostre que todo conjunto aberto contido em A está contido em $\mathrm{int}(A)$ e que todo conjunto fechado que contém A contém $\overline{A}$ (sugestão: use o resultado do Exercício B.1.17).

B.1.26. Exercício. Dado um espaço métrico M, mostre que:

(a) o conjunto vazio $\emptyset$ e o espaço M são ao mesmo tempo subconjuntos abertos e subconjuntos fechados de M;
(b) a união de uma família arbitrária de conjuntos abertos é um conjunto aberto;
(c) a interseção de dois conjuntos abertos (ou, mais geralmente, de uma família finita de conjuntos abertos) é um conjunto aberto;
(d) a interseção de uma família arbitrária de conjuntos fechados é um conjunto fechado;
(e) a união de dois conjuntos fechados (ou, mais geralmente, de uma família finita de conjuntos fechados) é um conjunto fechado.

(sugestão: para (d), (e) use (b), (c) e as chamadas *leis de De Morgan* que dizem que o complementar de uma união é igual à interseção dos complementares e que o complementar de uma interseção é igual à união dos complementares). Tomando $M = \mathbb{R}$ munido da métrica Euclideana, dê exemplos de famílias infinitas enumeráveis de conjuntos abertos cuja interseção não é um conjunto aberto e de famílias infinitas enumeráveis de conjuntos fechados cuja união não é um conjunto fechado.

B.1.27. Exercício. Seja M um espaço métrico. Se U é um conjunto aberto e F é um conjunto fechado, mostre que $U \setminus F$ é um conjunto aberto e que $F \setminus U$ é um conjunto fechado (sugestão: use o resultado do Exercício B.1.26 e as igualdades $U \setminus F = U \cap (M \setminus F)$, $F \setminus U = F \cap (M \setminus U)$).

B.1.28. Lema. *Toda bola aberta num espaço métrico M é um conjunto aberto e toda bola fechada é um conjunto fechado.*

Demonstração. Dados $x \in M$, $r > 0$ e $y \in \mathrm{B}(x, r)$ então segue facilmente da desigualdade triangular que a bola aberta $\mathrm{B}(y, s)$ está contida em $\mathrm{B}(x, r)$ para $s = r - d(x, y) > 0$. Logo $\mathrm{B}(x, r)$ é um conjunto aberto. Agora, se $y \in M \setminus \mathrm{B}[x, r]$, então $\mathrm{B}(y, s) \subset M \setminus \mathrm{B}[x, r]$ para $s = d(x, y) - r > 0$; de fato, a existência de um elemento z em $\mathrm{B}(y, s) \cap \mathrm{B}[x, r]$ nós daria as desigualdades:

$$d(x, y) \leq d(x, z) + d(z, y) < r + s = d(x, y).$$

Logo $M \setminus \mathrm{B}[x, r]$ é um conjunto aberto e portanto $\mathrm{B}[x, r]$ é fechado. $\square$

B.1.29. Corolário. *Se A é um subconjunto de um espaço métrico M então* $\mathrm{int}(A)$ *é aberto e* $\overline{A}$ *é fechado (tendo em conta o resultado do Exercício B.1.25, vê-se que* $\mathrm{int}(A)$ *é o* maior conjunto aberto *contido em* A *e* $\overline{A}$ *é o* menor conjunto fechado *que contém* A*).*

Demonstração. Dado $x \in \mathrm{int}(A)$, existe $r > 0$ tal que $\mathrm{B}(x, r) \subset A$; como $\mathrm{B}(x, r)$ é aberto, temos que $\mathrm{B}(x, r) \subset \mathrm{int}(A)$ e portanto x pertence ao interior de $\mathrm{int}(A)$. Logo $\mathrm{int}(A)$ é aberto. O fato que $\overline{A}$ é fechado segue então da primeira igualdade em (B.1.1). □

B.1.30. Exercício. Sejam A um subconjunto de um espaço métrico M e $x \in M$. Mostre que:

(a) x pertence ao interior de A se e somente se x pertence a um aberto contido em A;
(b) x pertence ao fecho de A (isto é, x é aderente a A) se e somente se todo conjunto aberto que contém x intercepta A;
(c) x pertence à fronteira de A se e somente se todo conjunto aberto que contém x intercepta A e $M \setminus A$.

B.1.31. Exercício. Mostre que um subconjunto A de um espaço métrico M é denso se e somente se todo conjunto aberto não vazio intercepta A.

B.1.32. Definição. Se M é um espaço métrico, dizemos que um subconjunto V de M é uma *vizinhança* de um ponto $x \in M$ se x pertence ao interior de V, isto é, (de acordo com o resultado do Exercício B.1.30) se x pertence a um aberto contido em V.

B.1.33. Exercício. Dados um subconjunto A de um espaço métrico M e um ponto $x \in M$, mostre que:

(a) x pertence ao interior de A se e somente se x possui uma vizinhança contida em A;
(b) x pertence ao fecho de A (isto é, x é aderente a A) se e somente se toda vizinhança de x intercepta A;
(c) x pertence à fronteira de A se e somente se toda vizinhança de x intercepta A e $M \setminus A$.

B.1.34. Proposição. *Sejam M um espaço métrico, A_1, …, A_n subconjuntos de M e $A = \bigcup_{i=1}^n A_i$. Temos que $\overline{A} = \bigcup_{i=1}^n \overline{A_i}$ (isto é,* o fecho de uma união finita é igual à união dos fechos*).*

Demonstração. Temos que $\overline{A}$ é fechado e contém A_i, donde $\overline{A}$ contém $\overline{A_i}$ para todo $i = 1, \ldots, n$ e portanto:

$$\bigcup_{i=1}^n \overline{A_i} \subset \overline{A}.$$

Além do mais, $\bigcup_{i=1}^n \overline{A_i}$ é um conjunto fechado (Exercício B.1.26) que contém A e portanto contém $\overline{A}$. Logo $\overline{A} = \bigcup_{i=1}^n \overline{A_i}$. □

B.1.35. Proposição. *Seja (M, d) um espaço métrico e N um subconjunto de M. Temos que um subconjunto A de N é aberto em $(N, d|_{N\times N})$ se e somente se existe um conjunto U aberto em (M, d) tal que $A = U \cap N$.*

Demonstração. Vamos usar a notação introduzida no enunciado do Exercício B.1.6 (bem como o resultado desse exercício). Se U é aberto em M então para todo $x \in U \cap N$ temos que existe $r > 0$ tal que $\mathrm{B}_M(x, r) \subset U$; daí:

$$\mathrm{B}_N(x, r) = \mathrm{B}_M(x, r) \cap N \subset U \cap N,$$

o que mostra que $A = U \cap N$ é aberto em N. Seja agora $A \subset N$ um conjunto aberto em N e vamos mostrar que existe um conjunto U aberto em M tal que $A = U \cap N$. Para cada $x \in A$, seja $r_x > 0$ tal que $\mathrm{B}_N(x, r_x) \subset A$. Defina:

$$U = \bigcup_{x\in A} \mathrm{B}_M(x, r_x).$$

Temos que U é aberto em M, sendo uma união de conjuntos abertos (veja Exercício B.1.26 e Lema B.1.28); como cada bola aberta contém seu centro, temos que $A \subset U$ e portanto $A \subset U \cap N$. Além do mais:

$$U \cap N = \bigcup_{x\in A} \big(\mathrm{B}_M(x, r_x) \cap N\big) = \bigcup_{x\in A} \mathrm{B}_N(x, r_x) \subset A,$$

donde $A = U \cap N$. □

B.1.36. Corolário. *Seja (M, d) um espaço métrico e N um subconjunto de M. Se $A \subset N$ é aberto em (M, d) então A é aberto em $(N, d|_{N\times N})$ e, reciprocamente, se A é aberto em $(N, d|_{N\times N})$* e *N é aberto em (M, d) então A é aberto em (M, d).*

Demonstração. Se A é aberto em M então $A = A \cap N$ é aberto em N; reciprocamente, se A é aberto em N e N é aberto em M então $A = U \cap N$, com U aberto em M e portanto A também é aberto em M, sendo interseção de dois abertos (Exercício B.1.26). □

B.1.37. Proposição. *Seja (M, d) um espaço métrico e N um subconjunto de M. Temos que um subconjunto B de N é fechado em $(N, d|_{N\times N})$ se e somente se existe um conjunto F fechado em (M, d) tal que $B = F \cap N$.*

Demonstração. Se F é fechado em M então $B = F \cap N$ é fechado em N pois, como:

$$N \setminus B = N \setminus F = (M \setminus F) \cap N,$$

a Proposição B.1.35 nos dá que $N \setminus B$ é aberto em N. Reciprocamente, se B é fechado em N então $N \setminus B$ é aberto em N e, pela Proposição B.1.35, existe U aberto em M com $N \setminus B = U \cap N$. Daí $B = (M \setminus U) \cap N$, onde $F = M \setminus U$ é fechado em M. □

B.1.38. Corolário. *Seja* (M, d) *um espaço métrico e* N *um subconjunto de* M. *Se* $B \subset N$ *é fechado em* (M, d) *então* B *é fechado em* $(N, d|_{N\times N})$ *e, reciprocamente, se* B *é fechado em* $(N, d|_{N\times N})$ e N *é fechado em* (M, d) *então* B *é fechado em* (M, d).

Demonstração. Se B é fechado em M então $B = B \cap N$ é fechado em N; reciprocamente, se B é fechado em N e N é fechado em M então $B = F \cap N$, com F fechado em M e portanto B também é fechado em M, sendo interseção de dois fechados (Exercício B.1.26). □

B.1.39. Definição. O *diâmetro* de um espaço métrico (M, d), denotado por $\operatorname{diam}(M, d)$ (ou, quando a métrica estiver subentendida pelo contexto, simplesmente por $\operatorname{diam}(M)$), é definido por:

$$\operatorname{diam}(M, d) = \sup_{x,y\in M} d(x, y) = \sup\big\{d(x, y) : x, y \in M\big\}. \tag{B.1.2}$$

Se A é um subconjunto de M então o *diâmetro* de A, denotado por $\operatorname{diam}(A)$, é definido como sendo o diâmetro do espaço métrico $(A, d|_{A\times A})$.

O supremo em (B.1.2) é entendido como um supremo na reta estendida[1] $\overline{\mathbb{R}} = [-\infty, +\infty]$; temos que esse supremo é igual a $+\infty$ se e somente se para todo $k \in \mathbb{R}$ existem $x, y \in M$ com $d(x, y) > k$ e temos que esse supremo é igual a $-\infty$ se e somente se M é vazio. Evidentemente, se M é não vazio então $\operatorname{diam}(M) \ge 0$.

Se A é um subconjunto de um espaço métrico M então o diâmetro de A é obviamente igual a:

$$\operatorname{diam}(A) = \sup_{x,y\in A} d(x, y).$$

B.1.40. Exercício. Sejam M um espaço métrico e A um subconjunto de M. Mostre que $\operatorname{diam}(A) \le \operatorname{diam}(M)$ (em particular, se $A \subset B \subset M$ então $\operatorname{diam}(A) \le \operatorname{diam}(B)$).

B.1.41. Exercício. Se M é um espaço métrico, mostre que para todo $x \in M$ e todo número real $r > 0$, o diâmetro da bola aberta $\mathrm{B}(x, r)$ e da bola fechada $\mathrm{B}[x, r]$ são menores ou iguais a $2r$.

B.1.42. Proposição. *Se* M *é um espaço métrico e* A *é um subconjunto de* M *então* $\operatorname{diam}(A) = \operatorname{diam}(\overline{A})$.

Demonstração. Temos $\operatorname{diam}(A) \le \operatorname{diam}(\overline{A})$, de modo que basta mostrar que $\operatorname{diam}(\overline{A}) \le \operatorname{diam}(A)$. Essa desigualdade é trivial se $\operatorname{diam}(A) = +\infty$,

[1] A reta real usual $\mathbb{R}$, junto com dois pontos adicionais $+\infty$, $-\infty$, munida da ordem total que coincide com a usual em $\mathbb{R}$ e tal que $-\infty < x < +\infty$, para todo $x \in \mathbb{R}$. Na reta estendida, *todo* conjunto possui supremo e ínfimo. Um conjunto ilimitado superiormente na reta real usual possui supremo igual a $+\infty$ em $\overline{\mathbb{R}}$ e o conjunto vazio possui supremo igual a $-\infty$ em $\overline{\mathbb{R}}$, já que qualquer elemento de $\overline{\mathbb{R}}$ é cota superior do vazio e $-\infty$ é então a menor cota superior do conjunto vazio.

de modo que podemos supor $\operatorname{diam}(A) < +\infty$. Sejam $x, y \in \overline{A}$. Para todo $\varepsilon > 0$, as bolas abertas $\mathrm{B}(x, \varepsilon)$, $\mathrm{B}(y, \varepsilon)$ interceptam A, isto é, existem $x', y' \in A$ com $d(x, x') < \varepsilon$ e $d(y, y') < \varepsilon$. Daí:

$$d(x, y) \le d(x, x') + d(x', y') + d(y', y) < \operatorname{diam}(A) + 2\varepsilon.$$

Como $\varepsilon > 0$ é arbitrário, concluímos que $d(x, y) \le \operatorname{diam}(A)$, para todos $x, y \in \overline{A}$. A conclusão segue. □

B.1.43. Definição. Um espaço métrico (M, d) é dito *limitado* se:

$$\operatorname{diam}(M) < +\infty.$$

Um subconjunto A de M é dito *limitado* quando o espaço métrico $(A, d|_{A\times A})$ é limitado.

Evidentemente, um subconjunto A de um espaço métrico M é limitado se e somente se existe $k \in \mathbb{R}$ tal que $d(x, y) \le k$, para todos $x, y \in A$.

B.1.44. Definição. Se X é um conjunto e M é um espaço métrico, uma função $f : X \to M$ é dita *limitada* se a sua imagem $f(X)$ é um subconjunto limitado de M.

B.1.45. Exercício. Mostre que um espaço métrico (M, d) é limitado se e somente se a função $d : M \times M \to \mathbb{R}$ é limitada, onde $\mathbb{R}$ é munido da métrica Euclideana.

B.1.46. Exercício. Seja M um espaço métrico. Mostre que:

(a) se M é limitado, todo subconjunto de M é limitado (em particular, um subconjunto de um subconjunto limitado é limitado);
(b) se $A, B \subset M$ são limitados então $A \cup B$ é limitado e, dados $x_0 \in A$, $y_0 \in B$ temos:

$$\operatorname{diam}(A \cup B) \le \operatorname{diam}(A) + \operatorname{diam}(B) + d(x_0, y_0);$$

(c) se A é um subconjunto limitado de M então para todo $x \in M$ existe $r > 0$ tal que $A \subset \mathrm{B}(x, r)$ (sugestão: se $A \ne \emptyset$, escolha $x_0 \in A$ e tome $r < d(x, x_0) + \operatorname{diam}(A)$).

B.1.47. Definição. Dados espaços métricos (M, d), (N, d'), uma função $f : M \to N$ é dita *Lipschitziana* se existe um número real $k \ge 0$ tal que $d'\big(f(x), f(y)\big) \le k\, d(x, y)$, para todos $x, y \in M$; um tal número real k é chamado uma *constante de Lipschitz* para f. Uma função Lipschitziana que admite uma constante de Lipschitz menor do que 1 é também chamada uma *contração* (quando 1 é uma constante de Lipschitz para f às vezes se diz que f é uma *contração fraca*).

Obviamente se k é uma constante de Lipschitz para f e $k \le k'$ então k' também é uma constante de Lipschitz para f.

B.1.48. Exercício. Mostre que a composição de aplicações Lipschitzianas é ainda Lipschitziana. Mais especificamente, se k_f é uma constante de Lipschitz para $f : M \to N$ e k_g é uma constante de Lipschitz para $g : N \to P$ então $k_f k_g$ é uma constante de Lipschitz para $g \circ f$.

B.1.49. Exemplo. Se $f : [a, b] \to \mathbb{R}$ é uma função contínua, derivável em $]a, b[$ e $|f'(x)| \leq k$, para todo $x \in \,]a, b[$ então segue do Teorema do Valor Médio que f é Lipschitziana e k é uma constante de Lipschitz para f.

B.1.50. Proposição. *Sejam (M, d), (N, d') espaços métricos com M limitado. Se existe uma função Lipschitziana sobrejetora $f : M \to N$ então N também é limitado. Além do mais, se M é não vazio e $k \geq 0$ é uma constante de Lipschitz para f então:*

$$\mathrm{diam}(N) \leq k \, \mathrm{diam}(M). \tag{B.1.3}$$

Demonstração. Se $M = \emptyset$ então $N = \emptyset$ e portanto N é limitado. Se $M \neq \emptyset$, provamos (B.1.3). Dados $x, y \in N$ existem $x_0, y_0 \in M$ com $x = f(x_0)$, $y = f(y_0)$ e portanto:

$$d'(x, y) \leq k \, d(x, y) \leq k \, \mathrm{diam}(M). \qquad \square$$

B.1.51. Corolário. *Se (M, d), (N, d') são espaços métricos então uma função Lipschitziana $f : M \to N$ leva subconjuntos limitados de M em subconjuntos limitados de N. Além do mais, se $A \subset M$ é não vazio e k é uma constante de Lipschitz para f então:*

$$\mathrm{diam}\big(f(A)\big) \leq k \, \mathrm{diam}(A).$$

Demonstração. Basta observar que $f|_A : A \to f(A)$ é uma função Lipschitziana (com constante de Lipschitz k) e sobrejetora. $\square$

B.1.52. Definição. Dados espaços métricos (M, d), (N, d') então uma aplicação $f : M \to N$ é dita uma *imersão isométrica* se:

$$d'\big(f(x), f(y)\big) = d(x, y),$$

para todos $x, y \in M$. Uma imersão isométrica sobrejetora é dita uma *isometria*.

Evidentemente toda imersão isométrica é Lipschitziana com constante de Lipschitz igual a 1.

B.1.53. Exercício. Mostre que:

(a) toda imersão isométrica é injetora;
(b) a aplicação inversa de uma isometria é ainda uma isometria;
(c) a composição de duas imersões isométricas (resp., isometrias) é ainda uma imersão isométrica (resp., isometria).

B.1.54. Exemplo. Se (M, d) é um espaço métrico e N é um subconjunto de M (munido da métrica induzida $d|_{N \times N}$, como sempre) então a aplicação inclusão $i : N \to M$ é uma imersão isométrica. Dados números naturais m, n com $m \leq n$ então a aplicação:

$$\mathbb{R}^m \ni (x_1, \ldots, x_m) \longmapsto (x_1, \ldots, x_m, 0, \ldots, 0) \in \mathbb{R}^n$$

é uma imersão isométrica.

B.1.55. Exercício. Se (M, d) é um espaço métrico, N é um conjunto e $f : N \to M$ é uma função injetora, mostre que existe *uma única* métrica d_f em N que torna f uma imersão isométrica (sugestão: d_f tem que ser definida por $d_f(x, y) = d\big(f(x), f(y)\big)$, para todos $x, y \in N$). A métrica d_f é às vezes chamada a *métrica induzida* em N por f e d (note que se $N \subset M$ e f é a aplicação inclusão então $d_f = d|_{N\times N}$).

B.2. Funções contínuas e uniformemente contínuas

B.2.1. Definição. Sejam (M, d), (N, d') espaços métricos. Uma função $f : M \to N$ é dita *contínua* num ponto $x \in M$ se para todo $\varepsilon > 0$ existe $\delta > 0$ tal que:

$$\text{(B.2.1)} \qquad \text{para todo } y \in M,\ d(y, x) < \delta \Longrightarrow d'\big(f(y), f(x)\big) < \varepsilon.$$

A função f é dita *contínua* se for contínua em todo ponto de M.

Obviamente a condição (B.2.1) é equivalente a $f\big(\mathrm{B}(x, \delta)\big) \subset \mathrm{B}\left(f(x), \varepsilon\right)$ ou a $\mathrm{B}(x, \delta) \subset f^{-1}\big(\mathrm{B}(f(x), \varepsilon)\big)$.

B.2.2. Exercício. Mostre que em (B.2.1) poderíamos trocar um ou ambos os sinais "$<$" por "$\leq$" e a noção de função contínua não seria alterada.

B.2.3. Exercício. Sobre uma função $f : M \to N$ e um ponto $x \in M$, mostre que são equivalentes as seguintes condições:

(i) f é contínua no ponto x;
(ii) para toda vizinhança V de $f(x)$ em N existe uma vizinhança U de x em M tal que $f(U) \subset V$;
(iii) para todo aberto V de N contendo $f(x)$ existe um aberto U de M contendo x tal que $f(U) \subset V$.

B.2.4. Exercício. Dados espaços métricos M, N, um subconjunto S de M, uma função $f : M \to N$ e um ponto $x \in S$, mostre que:

(a) se f é contínua no ponto x então a restrição $f|_S : S \to N$ também é contínua no ponto x;
(b) se $f|_S : S \to N$ é contínua no ponto x e x pertence ao interior de S então f é contínua no ponto x.

Dê um exemplo de uma situação em que $f|_S$ é contínua (isto é, contínua em todo ponto de S), mas f não é contínua em ponto algum de S (sugestão: tome $M = \mathbb{R}^2$, $S = \mathbb{R} \times \{0\}$ e $f : M \to \mathbb{R}$ uma função que vale 1 em S e 0 em $M \setminus S$).

B.2.5. Exercício. Dados espaços métricos M, N, um subconjunto S de N e uma função $f : M \to N$ tal que $f(M) \subset S$, considere a função $f_0 : M \to S$ que difere de f apenas pelo contra-domínio. Dado $x \in M$, mostre que f é contínua no ponto x se e somente se f_0 é contínua no ponto x, onde S é munido da métrica induzida da métrica de N.

B.2.6. Proposição. *Dados espaços métricos (M, d), (N, d'), (P, d'') e funções $f : M \to N$, $g : N \to P$, se f é contínua num ponto $x \in M$ e g é* *contínua no ponto $f(x)$ então a função composta $g \circ f$ é contínua no ponto x. Em particular, se f e g são contínuas então $g \circ f$ é contínua.*

Demonstração. Dado $\varepsilon > 0$ então, pela continuidade de g no ponto $f(x)$, existe $\rho > 0$ tal que:

$$\text{para todo } z \in N,\ d'\big(z, f(x)\big) < \rho \Longrightarrow d''\big[g(z), g\big(f(x)\big)\big] < \varepsilon; \tag{B.2.2}$$

a partir de ρ, a continuidade de f no ponto x nos dá $\delta > 0$ tal que $d(y, x) < \delta$ implica $d'\big(f(y), f(x)\big) < \rho$, para todo $y \in M$. Logo, se $y \in M$ e $d(y, x) < \delta$, podemos fazer $z = f(y)$ em (B.2.2) obtendo $d''\big((g\circ f)(y), (g\circ f)(x)\big) < \varepsilon$. □

B.2.7. Proposição. *Sejam M, N espaços métricos e $f : M \to N$ uma função. São equivalentes:*

(a) *f é contínua;*
(b) *para todo subconjunto aberto U de N, $f^{-1}(U)$ é aberto em M;*
(c) *para todo subconjunto fechado F de N, $f^{-1}(F)$ é fechado em M.*

Demonstração. A equivalência entre (b) e (c) decorre facilmente do fato que um conjunto é aberto se e somente se seu complementar é fechado e do fato que $f^{-1}(N \setminus U) = M \setminus f^{-1}(U)$. Vamos mostrar a equivalência entre (a) e (b). Suponha que f é contínua e que U é aberto em N e vamos mostrar que $f^{-1}(U)$ é aberto em M. Para todo ponto $x \in f^{-1}(U)$, temos $f(x) \in U$ e portanto existe $\varepsilon > 0$ tal que $\mathrm{B}\big(f(x), \varepsilon\big) \subset U$; daí, como f é contínua no ponto x, existe $\delta > 0$ tal que $f\big(\mathrm{B}(x, \delta)\big) \subset \mathrm{B}\big(f(x), \varepsilon\big)$ e portanto $\mathrm{B}(x, \delta) \subset f^{-1}(U)$. Logo $f^{-1}(U)$ é aberto. Agora assuma (b) e vamos mostrar que f é contínua. Sejam $x \in M$ e $\varepsilon > 0$ dados. Como $U = \mathrm{B}\big(f(x), \varepsilon\big)$ é aberto e contém $f(x)$, $f^{-1}(U)$ é aberto e contém x, donde existe $\delta > 0$ tal que $\mathrm{B}(x, \delta) \subset f^{-1}(U)$. Logo $f\big(\mathrm{B}(x, \delta)\big) \subset \mathrm{B}\big(f(x), \varepsilon\big)$, mostrando que f é contínua no ponto x. □

B.2.8. Definição. Sejam (M, d), (N, d') espaços métricos. Uma função $f : M \to N$ é dita um *homeomorfismo* se f é contínua, bijetora e se a função inversa $f^{-1} : N \to M$ também é contínua.

Segue diretamente da Proposição B.2.7 a seguinte:

B.2.9. Proposição. *Sejam M, N espaços métricos e $f : M \to N$ uma função bijetora. São equivalentes:*

(a) *f é um homeomorfismo;*
(b) *para todo subconjunto U de N, U é aberto em N se e somente se $f^{-1}(U)$ é aberto em M;*
(c) *para todo subconjunto F de N, F é fechado em N se e somente se $f^{-1}(F)$ é fechado em M.* □

B.2.10. Exercício. Mostre que a aplicação inversa de um homeomorfismo é ainda um homeomorfismo e que a composição de homeomorfismos é um homeomorfismo.

B.2.11. Definição. Sejam (M,d), (N,d') espaços métricos. Uma função $f : M \to N$ é dita *uniformemente contínua* se para todo $\varepsilon > 0$ existe $\delta > 0$ tal que, para todos $x, y \in M$, se $d(x,y) < \delta$ então $d\big(f(x), f(y)\big) < \varepsilon$. Uma função bijetora uniformemente contínua com inversa uniformemente contínua é dita um *homeomorfismo uniforme*.

Obviamente toda função uniformemente contínua é contínua e todo homeomorfismo uniforme é um homeomorfismo.

B.2.12. Exercício. Mostre que a composição de aplicações uniformemente contínuas (resp., de homeomorfismos uniformes) é uniformemente contínua (resp., um homeomorfismo uniforme).

B.2.13. Exemplo. Toda função constante é uniformemente contínua e a aplicação identidade de um espaço métrico é um homeomorfismo uniforme.

B.2.14. Exemplo. Toda função Lipschitziana é uniformemente contínua (e, em particular, contínua): de fato, se $k > 0$ é uma constante de Lipschitz para $f : M \to N$ então, para todo $\varepsilon > 0$, podemos tomar $\delta = \frac{\varepsilon}{k}$ na Definição B.2.11. Note que, em particular, toda imersão isométrica é uniformemente contínua e toda isometria é um homeomorfismo uniforme.

B.2.15. Exemplo. A função $s : [0, +\infty[\ \ni x \mapsto \sqrt{x} \in \mathbb{R}$ é uniformemente contínua. De fato, dado $\varepsilon > 0$, tome $\delta = \frac{\varepsilon^2}{2}$. Sejam $x, y \geq 0$ com $|x - y| < \delta$. Se $x \geq \frac{\varepsilon^2}{4}$ ou $y \geq \frac{\varepsilon^2}{4}$, temos $\sqrt{x} + \sqrt{y} \geq \frac{\varepsilon}{2}$ e portanto:

$$\big|\sqrt{x} - \sqrt{y}\,\big| = \frac{|x-y|}{\sqrt{x}+\sqrt{y}} \leq \frac{2}{\varepsilon}\,|x-y| < \frac{2\delta}{\varepsilon} = \varepsilon;$$

por outro lado, se $x < \frac{\varepsilon^2}{4}$ e $y < \frac{\varepsilon^2}{4}$, temos:

$$\big|\sqrt{x} - \sqrt{y}\,\big| \leq \sqrt{x} + \sqrt{y} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

No entanto, observamos que a função s não é Lipschitziana, pois se $k \geq 0$ fosse uma constante de Lipschitz para s teríamos $\sqrt{x} \leq kx$ para todo $x \geq 0$ e portanto $k \geq \frac{1}{\sqrt{x}}$, para todo $x > 0$. A aplicação $s : [0, +\infty[\ \to [0, +\infty[$ é também um exemplo de um homeomorfismo uniformemente contínuo que não é um homeomorfismo uniforme; de fato, a aplicação inversa

$$s^{-1} : [0, +\infty[\ \ni x \longmapsto x^2 \in [0, +\infty[$$

é contínua (deixamos isso a cargo do leitor), mas não é uniformemente contínua: se s^{-1} fosse uniformemente contínua, existiria $\delta > 0$ tal que $|x^2 - y^2| < 1$, para todos $x, y \geq 0$ com $|x - y| < \delta$. Mas fazendo $y = x + \frac{\delta}{2}$, obtemos $|x^2 - y^2| = \delta x + \frac{\delta^2}{4} > 1$, para $x > \frac{1}{\delta}$.

B.3. Limites de Funções

Antes de definirmos a noção de limite, precisamos da noção de ponto de acumulação.

B.3.1. Definição. Seja M um espaço métrico e seja A um subconjunto de M. Um ponto $x \in M$ é dito um *ponto de acumulação* de A se para todo $r > 0$ existe $y \in \mathrm{B}(x, r) \cap A$ com $y \neq x$ (isto é, o conjunto $\big(\mathrm{B}(x, r) \cap A\big) \setminus \{x\}$ não é vazio). O conjunto dos pontos de acumulação de A é chamado o *conjunto derivado* de A e é denotado por[2] A'.

B.3.2. Exercício. Sejam M um espaço métrico e $A \subset M$. Mostre que:

(a) x é ponto de acumulação de A se e somente se x é ponto aderente de $A \setminus \{x\}$, isto é:
$$x \in A' \Longleftrightarrow x \in \overline{A \setminus \{x\}};$$
(b) x é ponto de acumulação de A se e somente se x é ponto de acumulação de $A \setminus \{x\}$;
(c) todo ponto de acumulação de A é ponto aderente de A, isto é, $A' \subset \overline{A}$;
(d) se $x \in M \setminus A$ então x é ponto de acumulação de A se e somente se x é ponto aderente de A;
(e) $\overline{A} = A \cup A'$;
(f) A é fechado se e somente se $A' \subset A$.

B.3.3. Exercício. Se $M = \mathbb{R}$ e $A = \mathbb{Z}$, mostre que $\overline{A} = A$, mas $A' = \emptyset$.

B.3.4. Exercício. Sejam M um espaço métrico, $A \subset M$ e $x \in M$. Mostre que são equivalentes:

(a) x é ponto de acumulação de A;
(b) para todo aberto U em M, se $x \in U$ então $(U \cap A) \setminus \{x\} \neq \emptyset$;
(c) para toda vizinhança V de x temos $(V \cap A) \setminus \{x\} \neq \emptyset$.

B.3.5. Exercício. Sejam (M, d) um espaço métrico e N um subconjunto de M. Se A é um subconjunto de N e $x \in N$, mostre que x é ponto de acumulação de A relativamente ao espaço métrico $(N, d|_{N\times N})$ se e somente se x é ponto de acumulação de A relativamente ao espaço métrico (M, d) (sugestão: use o resultado do Exercício B.1.6). Conclua que o conjunto derivado de A relativamente ao espaço métrico $(N, d|_{N\times N})$ é igual à interseção com N do conjunto derivado de A relativamente ao espaço métrico (M, d).

B.3.6. Definição. Seja (M, d) um espaço métrico. Um ponto $x \in M$ que não é ponto de acumulação de M é dito um *ponto isolado* de M. Um espaço métrico no qual todos os seus pontos são isolados é chamado um espaço métrico *discreto*. Se A é um subconjunto de M então um ponto $x \in A$ é dito *isolado em* A se x é um ponto isolado no espaço métrico $(A, d|_{A\times A})$. O conjunto A é dito *discreto* se $(A, d|_{A\times A})$ é um espaço métrico discreto.

Evidentemente um subconjunto A de um espaço métrico M é discreto se e somente se todo ponto de A é isolado em A. Em vista do resultado do

[2]Embora, quando não há risco de confusão com o conjunto derivado, nós usaremos o apóstrofe também como parte do nome de um conjunto, como em "Sejam A, A' conjuntos…".

Exercício B.3.5, um ponto $x \in A$ é isolado em A se e somente se x não é ponto de acumulação de A relativamente ao espaço métrico (M, d).

B.3.7. Exercício. Seja M um espaço métrico. Mostre que:

(a) um ponto $x \in M$ é isolado se e somente se existe $r > 0$ tal que $\mathrm{B}(x, r) = \{x\}$;
(b) um ponto $x \in M$ é isolado se e somente se $\{x\}$ é um conjunto aberto em M;
(c) M é um espaço métrico discreto se e somente se todo subconjunto de M é aberto (sugestão: já que união de abertos é aberto, se os subconjuntos unitários são abertos então todo subconjunto é aberto);
(d) M é um espaço métrico discreto se e somente se todo subconjunto de M é fechado.

B.3.8. Exercício. Mostre que qualquer conjunto munido da métrica zero-um (Exemplo B.1.3) é um espaço métrico discreto.

B.3.9. Exercício. Sejam M um espaço métrico, A um subconjunto de M e $x \in A$. Mostre que:

(a) x é um ponto isolado de A se e somente se existe $r > 0$ tal que $\mathrm{B}(x, r) \cap A = \{x\}$;
(b) A é discreto se e somente se $A \cap A' = \emptyset$;
(c) A não possui pontos de acumulação em M se e somente se A é discreto *e* fechado.

B.3.10. Exercício. Se $M = \mathbb{R}$, mostre que o conjunto:

$$A = \left\{ \tfrac{1}{n} : n = 1, 2, \dots \right\}$$

é discreto, mas possui pontos de acumulação em M (a saber: $A' = \{0\}$).

B.3.11. Exercício. Sejam M, N espaços métricos e $f : M \to N$ uma função. Mostre que se $x \in M$ é um ponto isolado então f é contínua no ponto x. Conclua que se M é discreto então toda função $f : M \to N$ é contínua.

B.3.12. Proposição. *Sejam M um espaço métrico, A um subconjunto de M e $x \in M$ um ponto de acumulação de A. Para toda vizinhança V de x o conjunto $V \cap A$ é infinito.*

Demonstração. Se $V \cap A$ fosse finito, o conjunto $F = (V \cap A) \setminus \{x\}$ também seria finito e portanto fechado. Daí $U = \mathrm{int}(V) \setminus F$ seria um aberto que contém x e portanto, como x é ponto de acumulação de A, existiria $y \in U \cap A$ com $y \neq x$. Daí $y \in \mathrm{int}(V) \subset V$, $y \in A$ e $y \neq x$, donde $y \in F$, contradizendo $y \in U$. □

B.3.13. Definição. Sejam (M, d), (N, d') espaços métricos, A um subconjunto de M, $a \in M$ um ponto de acumulação de A e $f : A \to N$ uma função. Dizemos que $L \in N$ é um *limite* de f no ponto a se para todo $\varepsilon > 0$ dado existe $\delta > 0$ tal que:

$$\text{para todo } x \in A,\ 0 < d(x, a) < \delta \Longrightarrow d'\big(f(x), L\big) < \varepsilon. \tag{B.3.1}$$

Note que a condição (B.3.1) é equivalente a $f\big(\mathrm{B}(a,\delta)\setminus\{a\}\big)\subset\mathrm{B}(L,\varepsilon)$ ou a $\mathrm{B}(a,\delta)\setminus\{a\}\subset f^{-1}\big(\mathrm{B}(L,\varepsilon)\big)$.

O fato mais básico a ser estabelecido sobre limites é a seguinte:

B.3.14. Proposição. *Nas condições da Definição B.3.13, se um limite para f no ponto a existe ele é único.*

Demonstração. Sejam $L_1, L_2 \in N$ limites para f no ponto a. Para todo $\varepsilon > 0$, existem $\delta_1 > 0$, $\delta_2 > 0$ tais que:

$$\text{para todo } x \in A,\ 0 < d(x,a) < \delta_1 \Longrightarrow d'\big(f(x), L_1\big) < \varepsilon,$$
$$\text{para todo } x \in A,\ 0 < d(x,a) < \delta_2 \Longrightarrow d'\big(f(x), L_2\big) < \varepsilon.$$

Tomando $\delta = \min\{\delta_1,\delta_2\} > 0$, como a é ponto de acumulação de A, existe $x \in \mathrm{B}(a,\delta)\cap A$ diferente de a. Daí $d'\big(f(x),L_1\big) < \varepsilon$ e $d'\big(f(x),L_2\big) < \varepsilon$, donde, pela desigualdade triangular, $d'(L_1,L_2) < 2\varepsilon$. Como $\varepsilon > 0$ é arbitrário, temos $d'(L_1,L_2) = 0$ e portanto $L_1 = L_2$. □

Em vista da Proposição B.3.14, um limite para f no ponto a, quando existe, será chamado *o limite* de f no ponto a e será denotado por:

$$\lim_{x\to a} f(x).$$

B.3.15. Observação. Como queremos permitir que tome-se o limite de uma função f num ponto que não está em seu domínio, precisamos na Definição B.3.13 considerar um espaço métrico (M,d) que contém o domínio de f. Observamos que se trocássemos M por um subespaço de M que contém $A\cup\{a\}$ o limite de f no ponto a não seria alterado.

B.3.16. Exercício. Mostre que na Definição B.3.13 não faria diferença se trocássemos um ou ambos os sinais "$<$" por "$\le$" em (B.3.1).

B.3.17. Exercício. Sejam M, N espaços métricos, A um subconjunto de M, B um subconjunto de A, $a \in M$ um ponto de acumulação de B e $f : A \to N$ uma função. Dado $L \in N$, mostre que:

(a) se L é o limite de f no ponto a então L também é o limite de $f|_B$ no ponto a;

(b) se a possui uma vizinhança V em M tal que[3]:
$$(A\cap V)\setminus\{a\} = (B\cap V)\setminus\{a\}$$
e se L é o limite de $f|_B$ no ponto a então L é o limite de f no ponto a.

B.3.18. Exercício. Sejam M, N espaços métricos, A um subconjunto de M, B_1, B_2 subconjuntos de A e $f : A \to N$ uma função. Suponha que $a \in M$ é ponto de acumulação de B_1 e de B_2, que $L_1 \in N$ é o limite de $f|_{B_1}$ no ponto a e que $L_2 \in N$ é o limite de $f|_{B_2}$ no ponto a. Mostre que se $L_1 \ne L_2$ então f não possui limite no ponto a (sugestão: use o resultado do item (a) do Exercício B.3.17).

---

[3] Dois casos interessantes onde essa hipótese é satisfeita são: (i) se $a \in \mathrm{int}(B)$, podemos tomar $V = B$; (ii) se $B = A\setminus\{a\}$ podemos tomar $V = M$.

B.3.19. Exercício. Sejam M, N espaços métricos, A um subconjunto de M, $a \in M$ um ponto de acumulação de A, $f : A \to N$ uma função, S um subconjunto de M que contém $f(A)$ e $f_0 : A \to S$ a função que difere de f apenas pelo seu contra-domínio. Se $L \in S$ e S é munido da métrica induzida da métrica de N, mostre que L é o limite de f no ponto a se e somente se L é o limite de f_0 no ponto a.

B.3.20. Exercício. Sejam M, N espaços métricos, $f : M \to N$ uma função e $a \in M$ um ponto que não seja isolado (isto é, um ponto de acumulação de M). Mostre que f é contínua no ponto a se e somente se $f(a)$ é o limite de f no ponto a.

B.3.21. Exercício. Sejam M, N espaços métricos, A um subconjunto de M, $f : A \to N$ uma função e $a \in M$ um ponto de acumulação de A. Mostre que são equivalentes sobre $L \in N$:

(a) L é o limite de f no ponto a;
(b) para toda vizinhança V de L em N, existe uma vizinhança U de a em M tal que $f\big((U \cap A) \setminus \{a\}\big) \subset V$;
(c) para todo aberto V em N contendo L, existe um aberto U em M contendo a tal que $f\big((U \cap A) \setminus \{a\}\big) \subset V$.

B.3.22. Proposição. *Sejam (M, d), (N, d'), (P, d'') espaços métricos, A um subconjunto de M, $a \in M$ um ponto de acumulação de A e $f : A \to N$ uma função. Se existe e é igual a $L \in N$ o limite de f no ponto a e se $g : N \to P$ é uma função contínua no ponto L então existe o limite de $g \circ f$ no ponto a e é igual a $g(L)$; em símbolos:*

$$\lim_{x \to a} g\big(f(x)\big) = g\Big(\lim_{x \to a} f(x)\Big).$$

Demonstração. Seja dado $\varepsilon > 0$. A continuidade de g no ponto L nos dá $\rho > 0$ tal que $g\big(\mathrm{B}(L, \rho)\big) \subset \mathrm{B}\left(g(L), \varepsilon\right)$. Como $\lim_{x \to a} f(x) = L$, existe $\delta > 0$ tal que, para todo $x \in A$, se $0 < d(x, a) < \delta$ então $d'\big(f(x), L\big) < \rho$. Logo, se $x \in A$ e $0 < d(x, a) < \delta$ temos $f(x) \in \mathrm{B}(L, \rho)$ e portanto:

$$g\big(f(x)\big) \in \mathrm{B}\left(g(L), \varepsilon\right),$$

isto é, $d''\big((g \circ f)(x), g(L)\big) < \varepsilon$. □

B.3.23. Proposição. *Sejam (M, d), (N, d'), (P, d'') espaços métricos, A um subconjunto de M, $a \in M$ um ponto de acumulação de A, $f : A \to N$ uma função, B um subconjunto de N que contém $f(A)$ e $g : B \to P$ uma função. Suponha que o limite de f no ponto a existe e denote por $L \in N$ esse limite. Se a possui uma vizinhança V em M tal que $f(x) \neq L$, para todo $x \in (V \cap A) \setminus \{a\}$, então L é um ponto de acumulação de B em N; além do mais, se o limite de g no ponto L existe então o limite de $g \circ f : A \to P$ no ponto a existe e:*

$$\lim_{x \to a} g\big(f(x)\big) = \lim_{y \to L} g(y). \tag{B.3.2}$$

Demonstração. Para todo $\rho > 0$, como $\lim_{x\to a} f(x) = L$, existe $\delta > 0$ tal que $0 < d(x,a) < \delta$ implica $d'\big(f(x), L\big) < \rho$, para todo $x \in A$. Como a é ponto de acumulação de A e $\mathrm{B}(a,\delta) \cap V$ é uma vizinhança de a, existe $x \in \mathrm{B}(a,\delta) \cap V \cap A$ com $x \neq a$; daí $f(x) \in \mathrm{B}(L,\rho) \cap B$ e $f(x) \neq L$. Isso mostra que L é ponto de acumulação de B. Para mostrar (B.3.2), escreva $L' = \lim_{y\to L} g(y)$ e seja dado $\varepsilon > 0$; temos que existe $\rho > 0$ tal que:

$$\text{para todo } y \in B,\ 0 < d'(y,L) < \rho \Longrightarrow d''\big(g(y), L'\big) < \varepsilon. \tag{B.3.3}$$

Como $\lim_{x\to a} f(x) = L$, existe $\delta_1 > 0$ tal que $0 < d(x,a) < \delta_1$ implica $d'\big(f(x), L\big) < \rho$, para todo $x \in A$, e como V é uma vizinhança de a, existe também $\delta_2 > 0$ tal que $\mathrm{B}(a,\delta_2) \subset V$. Tome $\delta = \min\{\delta_1,\delta_2\} > 0$. Se $x \in A$ e $0 < d(x,a) < \delta$ então $d'\big(f(x), L\big) < \rho$; além do mais, como $x \in \mathrm{B}(a,\delta_2) \subset V$ e $x \neq a$, temos $f(x) \neq L$, donde podemos fazer $y = f(x)$ em (B.3.3) obtendo $d''\big((g\circ f)(x), L'\big) < \varepsilon$. □

B.4. Seqüências

B.4.1. Definição. Uma *seqüência* num conjunto M é uma função cujo domínio é o conjunto $\mathbb{N} = \{0,1,2,\ldots\}$ dos números naturais e o contra-domínio é M.

O valor no ponto $n \in \mathbb{N}$ de uma seqüência x é normalmente denotado por x_n em vez de $x(n)$ e a seqüência x é normalmente denotada por $(x_n)_{n\geq 0}$. Com pequeno abuso, nós também chamaremos de seqüência uma função cujo domínio é o conjunto $\mathbb{N}^* = \{1,2,\ldots\}$ dos números naturais não nulos; nesse caso, escrevemos $(x_n)_{n\geq 1}$ em vez de $(x_n)_{n\geq 0}$.

B.4.2. Definição. Sejam (M,d) um espaço métrico e $(x_n)_{n\geq 0}$ uma seqüência em M. Dizemos que $(x_n)_{n\geq 1}$ é *convergente* se existe $a \in M$ tal que para todo $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $d(x_n, a) < \varepsilon$, para todo $n \geq n_0$; um tal elemento $a \in M$ é chamado um *limite* para a seqüência $(x_n)_{n\geq 0}$.

B.4.3. Proposição. *Uma seqüência convergente possui um único limite.*

Demonstração. Se $a, a' \in M$ são ambos limites de uma seqüência $(x_n)_{n\geq 0}$ num espaço métrico M então para todo $\varepsilon > 0$, existem $n_1, n_2 \in \mathbb{N}$ tais que $d(x_n, a) < \varepsilon$, para todo $n \geq n_1$ e $d(x_n, a') < \varepsilon$, para todo $n \geq n_2$. Tomando $n = \max\{n_1, n_2\}$, vemos que $d(x_n, a) < \varepsilon$ e $d(x_n, a') < \varepsilon$, donde a desigualdade triangular nos dá $d(a,a') < 2\varepsilon$. Como $\varepsilon > 0$ é arbitrário, concluímos que $d(a,a') = 0$ e $a = a'$. □

Em vista da Proposição B.4.3, um limite para $(x_n)_{n\geq 0}$, quando existe, será chamado *o limite* de $(x_n)_{n\geq 0}$ e será denotado por:

$$\lim_{n\to\infty} x_n.$$

Se $a = \lim_{n\to\infty} x_n$ diremos também que $(x_n)_{n\geq 0}$ *converge* para a e escreveremos $x_n \to a$ ou, se houver possibilidade de confusão:

$$x_n \xrightarrow[n\to\infty]{} a.$$

B.4.4. Exemplo. Se M é um espaço métrico arbitrário e $a \in M$, então a *seqüência constante* definida por $x_n = a$ para todo $n \in \mathbb{N}$ converge para a.

B.4.5. Definição. Uma seqüência $(x_n)_{n\geq 0}$ num espaço métrico M é dita *limitada* quando o conjunto $\{x_n : n \in \mathbb{N}\}$ for limitado.

B.4.6. Proposição. *Toda seqüência convergente num espaço métrico M é limitada.*

Demonstração. Se $(x_n)_{n\geq 0}$ converge para a então existe $n_0 \in \mathbb{N}$ tal que $x_n \in \mathrm{B}(a,1)$, para todo $n \geq n_0$. Daí o conjunto $\{x_n : n \geq n_0\}$ é limitado. Mas o conjunto $\{x_n : n < n_0\}$ é finito, donde limitado, e daí também o conjunto $\{x_n : n \in \mathbb{N}\}$ é limitado, sendo união de dois conjuntos limitados. □

B.4.7. Definição. Uma seqüência $(x_n)_{n\geq 0}$ em $\mathbb{R}$ é dita *crescente* (resp., *decrescente*) se $x_n \leq x_m$ (resp., $x_n \geq x_m$), para quaisquer $n, m \in \mathbb{N}$ tais que $n \leq m$. A seqüência $(x_n)_{n\geq 0}$ é dita *estritamente crescente* (resp., *estritamente decrescente*) se $x_n < x_m$ (resp., $x_n > x_m$) para quaisquer $n, m \in \mathbb{N}$ tais que $n < m$. Uma seqüência que é ou crescente ou decrescente é dita *monótona*.

B.4.8. Exemplo. Toda seqüência monótona limitada em $\mathbb{R}$ é convergente. De fato, se $(x_n)_{n\geq 0}$ é crescente e limitada então $s = \sup\{x_n : n \in \mathbb{N}\}$ é o limite de $(x_n)_{n\geq 0}$, pois para todo $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ com $x_{n_0} > s - \varepsilon$ e daí temos $s - \varepsilon < x_{n_0} \leq x_n \leq s$, para todo $n \geq n_0$. Analogamente, se $(x_n)_{n\geq 0}$ é decrescente e limitada então $\inf\{x_n : n \in \mathbb{N}\}$ é o limite de $(x_n)_{n\geq 0}$.

B.4.9. Exercício. Sejam (M, d) um espaço métrico, N um subconjunto de M, $(x_n)_{n\geq 0}$ uma seqüência em N e a um ponto de N. Mostre que $(x_n)_{n\geq 0}$ converge para a no espaço métrico (M, d) se e somente se $(x_n)_{n\geq 0}$ converge para a no espaço métrico $(N, d|_{N\times N})$. Conclua que se $(x_n)_{n\geq 0}$ é convergente em (M, d) e o limite de $(x_n)_{n\geq 0}$ não está em N então $(x_n)_{n\geq 0}$ não é convergente em $(N, d|_{N\times N})$.

As seguintes terminologias são convenientes quando lidamos com seqüências: dada uma certa propriedade $\mathfrak{P}$ relativa a números naturais, dizemos que $\mathfrak{P}$ vale *para todo n suficientemente grande* se existe $n_0 \in \mathbb{N}$ tal que todo $n \geq n_0$ satisfaz $\mathfrak{P}$ (em outras palavras, o conjunto dos $n \in \mathbb{N}$ que não satisfazem $\mathfrak{P}$ é finito). Dizemos que $\mathfrak{P}$ *vale para todo n arbitrariamente grande* se para todo $n_0 \in \mathbb{N}$ existe $n \geq n_0$ que satisfaz $\mathfrak{P}$ (em outras palavras, o conjunto dos $n \in \mathbb{N}$ que satisfazem $\mathfrak{P}$ é infinito). Assim, por exemplo, temos que uma seqüência $(x_n)_{n\geq 0}$ converge para $a \in M$ se e somente se para todo $\varepsilon > 0$ a desigualdade $d(x_n, a) < \varepsilon$ vale para todo n suficientemente grande.

B.4.10. Exercício. Sejam M um espaço métrico, $(x_n)_{n\geq 0}$ uma seqüência em M e $a \in M$. Mostre que são equivalentes:

(a) $\lim_{n\to\infty} x_n = a$;

(b) para toda vizinhança V de a, temos $x_n \in V$, para todo n suficientemente grande;
(c) para todo conjunto aberto U em M com $a \in U$, temos $x_n \in U$, para todo n suficientemente grande;
(d) a seqüência $\big(d(x_n, a)\big)_{n\geq 0}$ converge para 0 no espaço métrico $\mathbb{R}$ (munido da métrica Euclideana).

B.4.11. Exercício. Sejam M um espaço métrico, $(x_n)_{n\geq 0}$, $(y_n)_{n\geq 0}$ seqüências em M e $a \in M$ um ponto. Se o conjunto $\big\{n \in \mathbb{N} : x_n \neq y_n\big\}$ é finito, mostre que $(x_n)_{n\geq 0}$ converge para a se e somente se $(y_n)_{n\geq 0}$ converge para a.

B.4.12. Definição. Dada uma seqüência $x : \mathbb{N} \to M$ num conjunto M então uma *subseqüência* de x é uma seqüência da forma $x \circ \sigma$, onde $\sigma : \mathbb{N} \to \mathbb{N}$ é uma função estritamente crescrente (isto é, $k < l$ implica $\sigma(k) < \sigma(l)$).

Usualmente, nós omitimos o nome da função σ e denotâmo-la simplesmente por $\mathbb{N} \ni k \mapsto n_k \in \mathbb{N}$; a subseqüência $x \circ \sigma$ de $x = (x_n)_{n\geq 0}$ é então denotada por $(x_{n_k})_{k\geq 0}$ (ou $(x_{n_k})_{k\geq 1}$, se o domínio de σ for $\mathbb{N}^*$).

B.4.13. Proposição. *Se uma seqüência $(x_n)_{n\geq 0}$ num espaço métrico M converge para $a \in M$ então qualquer subseqüência $(x_{n_k})_{k\geq 0}$ converge para a.*

Demonstração. Dado $\varepsilon > 0$, existe $m \in \mathbb{N}$ tal que $d(x_n, a) < \varepsilon$, para todo $n \geq m$. Como a função $k \mapsto n_k$ é estritamente crescente, existe k_0 tal que $n_k \geq m$, para todo $k \geq k_0$. Daí $d(x_{n_k}, a) < \varepsilon$, para todo $k \geq k_0$. □

B.4.14. Definição. Sejam M um espaço métrico e $(x_n)_{n\geq 0}$ uma seqüência em M. Dizemos que $a \in M$ é um *valor de aderência* para $(x_n)_{n\geq 0}$ quando existe uma subseqüência de $(x_n)_{n\geq 0}$ que converge para a.

Segue da Proposição B.4.13 que se uma seqüência é convergente então seu limite é o seu único valor de aderência. Assim, uma seqüência que possui dois valores de aderência distintos não pode ser convergente.

B.4.15. Exercício. Sejam (M, d) um espaço métrico, N um subconjunto de M, $(x_n)_{n\geq 0}$ uma seqüência em N e $a \in N$ um ponto. Mostre que a é valor de aderência de $(x_n)_{n\geq 0}$ no espaço métrico (M, d) se e somente se a é valor de aderência de $(x_n)_{n\geq 0}$ no espaço métrico $(N, d|_{N\times N})$.

B.4.16. Exercício. Considere a seqüência $(x_n)_{n\geq 0}$ em $\mathbb{R}$ definida por $x_n = 0$, se n é par, e $x_n = n$, se n é ímpar. Mostre que 0 é o único valor de aderência de $(x_n)_{n\geq 0}$, mas que $(x_n)_{n\geq 0}$ não é convergente.

B.4.17. Proposição. *Sejam M um espaço métrico e $(x_n)_{n\geq 0}$ uma seqüência em M. Temos que $a \in M$ é um valor de aderência para $(x_n)_{n\geq 0}$ se e somente se para todo aberto U contendo a temos que $x_n \in U$ para todo n arbitrariamente grande, isto é, se e somente se para todo aberto U contendo a o conjunto:*

$$\big\{n \in \mathbb{N} : x_n \in U\big\} \tag{B.4.1}$$

é infinito.

Demonstração. Suponha que a é um valor de aderência para $(x_n)_{n\geq 0}$ e seja $(x_{n_k})_{k\geq 0}$ uma subseqüência de $(x_n)_{n\geq 0}$ que converge para a. Dado um conjunto aberto U contendo a, temos que existe $k_0 \in \mathbb{N}$ tal que $x_{n_k} \in U$, para todo $k \geq k_0$; dai o conjunto (B.4.1) contém $\{n_k : k \geq k_0\}$ e é portanto infinito. Reciprocamente, suponha que para todo aberto U contendo a o conjunto (B.4.1) é infinito. Vamos construir recursivamente uma função estritamente crescente $k \mapsto n_k$ de modo que $(x_{n_k})_{k\geq 1}$ convirja para a. Em primeiro lugar[4], seja $n_1 \in \mathbb{N}$ tal que x_{n_1} pertence ao aberto $\mathrm{B}(a, 1)$. Assumindo n_k definido, seja $n_{k+1} \in \mathbb{N}$ tal que $n_{k+1} > n_k$ e $x_{n_{k+1}}$ pertence ao aberto $\mathrm{B}\left(a, \frac{1}{k+1}\right)$; note que tal n_{k+1} existe já que o conjunto:

$$\left\{n \in \mathbb{N} : x_n \in \mathrm{B}\left(a, \tfrac{1}{k+1}\right)\right\}$$

é infinito. Obtemos então uma subseqüência $(x_{n_k})_{k\geq 1}$ de $(x_n)_{n\geq 0}$ tal que $d(x_{n_k}, a) < \frac{1}{k}$, para todo $k \geq 1$, donde segue que $\lim_{k\to\infty} x_{n_k} = a$. □

Vários dos conceitos estudados nas seções anteriores podem ser caracterizados usando a noção de limite de seqüência, como veremos a seguir.

B.4.18. Proposição. *Sejam M um espaço métrico e A um subconjunto de M. Temos que um ponto $a \in M$ é interior a A se e somente se para toda seqüência $(x_n)_{n\geq 0}$ em M tal que $x_n \to a$ temos $x_n \in M$, para todo n suficientemente grande. Em particular, A é aberto se e somente se para toda seqüência convergente $(x_n)_{n\geq 0}$ em M com $\lim_{n\to\infty} x_n \in A$ temos $x_n \in A$, para todo n suficientemente grande.*

Demonstração. Se $a \in \mathrm{int}(A)$ e $x_n \to a$ então A é uma vizinhança de a e portanto $x_n \in A$, para todo n suficientemente grande (Exercício B.4.10). Reciprocamente, suponha que *para toda* seqüência $(x_n)_{n\geq 0}$ com $x_n \to a$ vale que $x_n \in A$, para todo n suficientemente grande. Se a não fosse um ponto interior a A, então para todo número natural $n \geq 1$ não seria o caso que $\mathrm{B}\left(a, \frac{1}{n}\right) \subset A$ e portanto existiria $x_n \in \mathrm{B}\left(a, \frac{1}{n}\right)$ tal que $x_n \notin A$. Daí $d(x_n, a) < \frac{1}{n} \to 0$ e portanto $(x_n)_{n\geq 1}$ converge para a. Mas não existe nenhum n com $x_n \in A$, contradizendo nossas hipóteses. □

B.4.19. Corolário. *Sejam M um espaço métrico e A um subconjunto de M. Temos que $a \in M$ é um ponto aderente a A (isto é, $a \in \overline{A}$) se e somente se existe uma seqüência $(x_n)_{n\geq 0}$ em A tal que $x_n \to a$. Em particular, A é fechado se e somente se para todo seqüência $(x_n)_{n\geq 0}$ em A que seja convergente em M temos $\lim_{n\to\infty} x_n \in A$.*

Demonstração. Se $x_n \in A$ para todo n e $x_n \to a$ então não pode ser o caso que a pertence ao interior de $M \setminus A$, pois caso contrário teríamos que $x_n \in M \setminus A$, para todo n suficientemente grande. Logo a pertence a $M\setminus \mathrm{int}(M\setminus A)$, que é igual ao fecho de A (Exercício B.1.13). Reciprocamente,

[4]Se fazemos mesmo questão que as seqüências estejam indexadas em $\mathbb{N} = \{0, 1, 2, \ldots\}$ podemos obviamente colocar um valor qualquer para n_0, digamos $n_0 = 0$, e depois tomar o cuidado de escolher $n_1 > n_0$.

se a é aderente a A então a não pertence ao interior de $M \setminus A$ e portanto existe uma seqüência $(x_n)_{n \geq 0}$ em M com $x_n \to a$ e tal que não é o caso que $x_n \in M \setminus A$ para todo n suficientemente grande; em outras palavras, $x_n \in A$ para todo n arbitrariamente grande. Existe então uma subseqüência $(x_{n_k})_{k \geq 0}$ de $(x_n)_{n \geq 0}$ (que converge para a e) tal que $x_{n_k} \in A$, para todo k. □

B.4.20. COROLÁRIO. *Sejam M um espaço métrico e A um subconjunto de M. Temos que $a \in M$ é um ponto de acumulação de A se e somente se existe uma seqüência $(x_n)_{n \geq 0}$ em $A \setminus \{a\}$ tal que $x_n \to a$.*

DEMONSTRAÇÃO. Segue do Corolário B.4.19 e do resultado do item (a) do Exercício B.3.2. □

B.4.21. PROPOSIÇÃO. *Sejam (M, d), (N, d') espaços métricos, A um subconjunto de M, $f : A \to N$ uma função e $a \in M$ um ponto de acumulação de A. Dado $L \in N$ então são equivalentes:*

(a) $\lim_{x \to a} f(x) = L$;
(b) *para toda seqüência $(x_n)_{n \geq 0}$ em $A \setminus \{a\}$ com $x_n \to a$ vale que $f(x_n) \to L$.*

DEMONSTRAÇÃO. Assuma (a) e seja $(x_n)_{n \geq 0}$ uma seqüência em $A \setminus \{a\}$ que converge para a. Dado $\varepsilon > 0$, como $\lim_{x \to a} f(x) = L$, existe $\delta > 0$ tal que $0 < d(x, a) < \delta$ implica $d'\big(f(x), L\big) < \varepsilon$, para todo $x \in A$. Como $x_n \to a$, existe $n_0 \in \mathbb{N}$ tal que $d(x_n, a) < \delta$, para todo $n \geq n_0$. Como $x_n \neq a$ para todo n, temos $0 < d(x_n, a) < \delta$, para todo $n \geq n_0$ e portanto:

$$d'\big(f(x_n), L\big) < \varepsilon.$$

Agora assuma (b). Se não fosse o caso que (a), existiria $\varepsilon > 0$ tal que para todo $\delta > 0$ existe $x \in A$ com $0 < d(x, a) < \delta$ e $d'\big(f(x), L\big) \geq \varepsilon$. Para cada $n \geq 1$, seja $x_n \in A$ com $0 < d(x_n, a) < \frac{1}{n}$ e $d'\big(f(x_n), L\big) \geq \varepsilon$. Daí $(x_n)_{n \geq 1}$ é uma seqüência em $A \setminus \{a\}$, $x_n \to a$ e não é o caso que $f(x_n) \to L$, contradizendo (b). □

B.4.22. PROPOSIÇÃO. *Dados espaços métricos (M, d), (N, d'), uma função $f : M \to N$ e $a \in M$, são equivalentes:*

(a) *f é contínua no ponto a;*
(b) *para qualquer seqüência $(x_n)_{n \geq 0}$ em M com $x_n \to a$ vale que $f(x_n) \to f(a)$.*

DEMONSTRAÇÃO. Assuma (a) e seja $(x_n)_{n \geq 0}$ uma seqüência em M que converge para a. Dado $\varepsilon > 0$, existe $\delta > 0$ tal que $d(x, a) < \delta$ implica $d'\big(f(x), f(a)\big) < \varepsilon$, para todo $x \in M$. Como $x_n \to a$, existe $n_0 \in \mathbb{N}$ tal que $d(x_n, a) < \delta$, para todo $n \geq n_0$. Daí $d'\big(f(x_n), f(a)\big) < \varepsilon$, para todo $n \geq n_0$. Agora assuma (b). Se não fosse o caso que (a), existiria $\varepsilon > 0$ tal que para todo $\delta > 0$ existe $x \in M$ com $d(x, a) < \delta$ e $d'\big(f(x), L\big) \geq \varepsilon$. Para cada $n \geq 1$, seja $x_n \in M$ com $d(x_n, a) < \frac{1}{n}$ e $d'\big(f(x_n), L\big) \geq \varepsilon$. Daí

$(x_n)_{n\geq 1}$ é uma seqüência em M, $x_n \to a$ e não é o caso que $f(x_n) \to f(a)$, contradizendo (b). $\square$

B.5. Topologia e equivalência entre métricas

B.5.1. DEFINIÇÃO. A *topologia* de um espaço métrico (M, d) é o conjunto $\tau(M, d)$ de todos os subconjuntos abertos de M:

$$\tau(M, d) = \big\{U \in \wp(M) : \text{para todo } x \in U, \text{ existe } r > 0 \text{ com } \mathrm{B}(x, r) \subset U\big\},$$

onde $\wp(M)$ denota o conjunto de todas as partes de M.

Embora a métrica d não apareça explicitamente na definição de $\tau(M, d)$ acima, é evidente que $\tau(M, d)$ pode depender (e normalmente depende) de d, já que o uso da notação de bola aberta $\mathrm{B}(x, r)$ faz uma referência implícita à métrica d. No entanto, por simplicidade, escreveremos às vezes apenas $\tau(M)$ quando d estiver subentendida pelo contexto e não houver risco de confusão.

Vários dos conceitos estudados até agora neste apêndice podem ser definidos fazendo referência apenas à topologia dos espaços métricos envolvidos, sem que seja feita referência direta às métricas. Um conceito que pode ser descrito fazendo referência apenas às topologias, sem menção direta[5] às métricas é chamado um *conceito topológico*. Por exemplo, a Proposição B.2.7 nos diz que o conceito de função contínua é topológico: uma função $f : M \to N$ é contínua se e somente se $f^{-1}(U) \in \tau(M)$, para todo $U \in \tau(N)$. Mais geralmente, o resultado do Exercício B.2.3 nos garante que também o conceito de continuidade num ponto é topológico: $f : M \to N$ é contínua num ponto $x \in M$ se e somente se para todo $V \in \tau(N)$ com $f(x) \in V$ existe $U \in \tau(M)$ com $x \in U$ e $f(U) \subset V$. Todos os conceitos introduzidos na Definição B.1.7 são topológicos (veja Exercício B.1.30). Os conceitos de conjunto aberto e fechado são (trivialmente) topológicos. Também são topológicos os conceitos de conjunto denso, de vizinhança de um ponto, de homeomorfismo, de ponto de acumulação (Exercício B.3.4), de ponto isolado, de espaço discreto, de limite de função (Exercício B.3.21) e de limite de seqüência (Exercício B.4.10). Veremos que os conceitos de função uniformemente contínua e de conjunto limitado não são topológicos (veja Observação B.5.6 e Exemplo B.5.13 abaixo).

B.5.2. DEFINIÇÃO. Duas métricas d_1, d_2 num conjunto M são ditas *equivalentes* (ou *topologicamente equivalentes*) se $\tau(M, d_1) = \tau(M, d_2)$, isto é, se os espaços métricos (M, d_1) e (M, d_2) possuem os mesmos conjuntos abertos.

Temos então que os conceitos topológicos são aqueles que permanecem invariantes quando trocamos uma métrica por outra equivalente. Por exemplo, se $f : M \to N$ é uma função contínua quando M, N são munidos

[5]Evidentemente, uma referência à topologia é uma referência indireta à métrica, já que a topologia depende da métrica.

respectivamente de métricas d, d' então f ainda será contínua se trocarmos d e d' respectivamente por métricas equivalentes d_1 e d'_1.

B.5.3. PROPOSIÇÃO. *Duas métricas d_1, d_2 num conjunto M são equivalentes se e somente se a aplicação identidade de (M, d_1) para (M, d_2) é um homeomorfismo.*

DEMONSTRAÇÃO. Segue diretamente da Proposição B.2.9. □

B.5.4. DEFINIÇÃO. Duas métricas d_1, d_2 num conjunto M são ditas *uniformemente equivalentes* se a aplicação identidade de (M, d_1) para (M, d_2) é um homeomorfismo uniforme (Definição B.2.11). As métricas d_1, d_2 são ditas *Lipschitz-equivalentes* se a aplicação identidade de (M, d_1) para (M, d_2) e a aplicação identidade de (M, d_2) para (M, d_1) são Lipschitzianas.

Como toda aplicação Lipschitziana é uniformemente contínua, temos que métricas Lipschitz-equivalentes são uniformemente equivalentes e como toda aplicação uniformemente contínua é contínua, temos que duas métricas uniformemente equivalentes são equivalentes.

B.5.5. EXERCÍCIO. Mostre que se d e d' são métricas Lipschitz-equivalentes em M então (M, d) é limitado se e somente se (M, d') é limitado (sugestão: use a Proposição B.1.50).

B.5.6. OBSERVAÇÃO. No Exercício B.4 pedimos ao leitor para demonstrar que toda métrica é uniformemente equivalente a uma métrica limitada. Vemos então que não se pode trocar "Lipschitz-equivalentes" por "uniformemente equivalentes" no enunciado do Exercício B.5.5 (e, em particular, vemos que "Lipschitz-equivalente" não é o mesmo que "uniformemente equivalente"). Vemos também que o conceito de espaço métrico limitado não é topológico: podemos ter duas métricas equivalentes (e até uniformemente equivalentes), sendo uma limitada e a outra não.

B.5.7. EXERCÍCIO. Sejam d_1, d_2 métricas num conjunto M. Mostre que:

(a) d_1 e d_2 são uniformemente equivalentes se e somente se para todo $\varepsilon > 0$ existe $\delta > 0$ tal que, para todos $x, y \in M$, $d_1(x, y) < \delta$ implica $d_2(x, y) < \varepsilon$ e $d_2(x, y) < \delta$ implica $d_1(x, y) < \varepsilon$;

(b) d_1 e d_2 são Lipschitz-equivalentes se e somente se existem números reais $k, k' > 0$ tais que:

$$k'\, d_1(x, y) \le d_2(x, y) \le k\, d_1(x, y), \quad \text{para todos } x, y \in M. \tag{B.5.1}$$

B.5.8. EXERCÍCIO. Mostre que, dado um conjunto M, as relações "ser equivalente a", "ser uniformemente equivalente a" e "ser Lipschitz-equivalente a" são relações de equivalência (isto é, são relações reflexivas, simétricas e transitivas) no conjunto de todas as métricas em M (sugestão: use os resultados dos Exercícios B.2.10, B.2.12 e B.1.48).

B.5.9. EXERCÍCIO. Sejam (M, d), (N, d') espaços métricos e d_1, d'_1 métricas uniformemente equivalentes a d e a d', respectivamente. Dada uma função $f : M \to N$, mostre que $f : (M, d) \to (N, d')$ é uniformemente contínua se e somente se a aplicação $f : (M, d_1) \to (N, d_1')$ é uniformemente contínua (sugestão: use o resultado do Exercício B.2.12).

B.5.10. Exercício. Sejam (M, d), (N, d') espaços métricos e d_1, d_1' métricas Lipschitz-equivalentes a d e a d', respectivamente. Dada uma função $f : M \to N$, mostre que $f : (M, d) \to (N, d')$ é Lipschitziana se e somente se a aplicação $f : (M, d_1) \to (N, d_1')$ é Lipschitziana (sugestão: use o resultado do Exercício B.1.48).

B.5.11. Exemplo. As métricas d, d_1 e d_∞ em $\mathbb{R}^n$ definidas no Exemplo B.1.2 são duas a duas Lipschitz-equivalentes. De fato, é fácil ver que:

$$d_\infty(x, y) \le d_1(x, y) \le n\, d_\infty(x, y), \quad d_\infty(x, y) \le d(x, y) \le \sqrt{n}\, d_\infty(x, y),$$

para todos $x, y \in \mathbb{R}^n$, donde d_∞ é Lipschitz-equivalente a d_1 e a d. Por transitividade, vemos que d é Lipschitz-equivalente a d_1.

B.5.12. Proposição. *Sejam (M, d), (N, d') espaços métricos e seja dada uma função bijetora $f : M \to N$; denote por d_f a métrica em M induzida por f e por d' (veja Exercício B.1.55). Temos que:*

(a) *d_f é equivalente a d se e somente se a aplicação $f : (M, d) \to (N, d')$ é um homeomorfismo;*
(b) *d_f é uniformemente equivalente a d se e somente se a aplicação $f : (M, d) \to (N, d')$ é um homeomorfismo uniforme;*
(c) *d_f é Lipschitz equivalente e d se e somente se ambas as aplicações $f : (M, d) \to (N, d')$, $f^{-1} : (N, d') \to (M, d)$ são Lipschitzianas.*

Demonstração. Note que a aplicação $f : (M, d) \to (N, d')$ é igual à composição da aplicação identidade $\mathrm{Id} : (M, d) \to (M, d_f)$ com a isometria $f : (M, d_f) \to (N, d')$; em outras, palavras o diagrama:

$$\begin{tikzcd}
(M, d) \arrow[r, "f"] \arrow[d, "\mathrm{Id}"'] & (N, d') \\
(M, d_f) \arrow[ur, "f \text{ (isometria)}"'] &
\end{tikzcd}$$

é comutativo. Já que uma isometria é Lipschitziana (e portanto uniformemente contínua e contínua), segue facilmente que $f : (M, d) \to (N, d')$ tem as mesmas propriedades (continuidade, continuidade uniforme ou existência de constante de Lipschitz) que a aplicação identidade $\mathrm{Id} : (M, d) \to (M, d_f)$. Uma observação completamente análoga vale para $f^{-1} : (N, d') \to (M, d)$ e $\mathrm{Id}^{-1} : (M, d_f) \to (M, d)$. $\square$

B.5.13. Exemplo. Seja $M = [0, +\infty[$ e denote por d a restrição a M da métrica Euclideana de $\mathbb{R}$. A função $f : M \ni x \mapsto x^2 \in M$ é um homeomorfismo de (M, d) em (M, d) e portanto, pela Proposição B.5.12, a métrica $d_f : M \times M \ni (x, y) \mapsto |x^2 - y^2|$ é equivalente a d. Temos então que a função $f : (M, d_f) \to (M, d)$ é uniformemente contínua (pois é uma isometria), mas a função $f : (M, d) \to (M, d)$ não é (veja Exemplo B.2.15). Em particular, continuidade uniforme não é um conceito topológico.

B.6. Métricas no produto cartesiano

Dados espaços métricos (M_1, d_1), ..., (M_n, d_n), há muitas maneiras naturais de se definir uma métrica no produto cartesiano $M = \prod_{i=1}^n M_i$. Três delas são:

$$D_1 : M \times M \ni (x, y) \longmapsto \sum_{i=1}^n d_i(x_i, y_i) \in \mathbb{R}, \tag{B.6.1}$$

$$D_2 : M \times M \ni (x, y) \longmapsto \Big(\sum_{i=1}^n d_i(x_i, y_i)^2\Big)^{\frac{1}{2}} \in \mathbb{R}, \tag{B.6.2}$$

$$D_\infty : M \times M \ni (x, y) \longmapsto \max\big\{d_i(x_i, y_i) : i = 1, \dots, n\big\} \in \mathbb{R}. \tag{B.6.3}$$

Deixamos a cargo do leitor a verificação do fato que D_1, D_2 e D_∞ são mesmo métricas (veja Exercício B.5). As métricas D_1, D_2 e D_∞ são chamadas respectivamente a *métrica produto tipo* 1, a *métrica produto tipo* 2 e a *métrica produto tipo* ∞ associadas às métricas d_i, $i = 1, \dots, n$.

As desigualdades (compare com o Exemplo B.5.11):

$$\begin{aligned} D_\infty(x, y) &\le D_1(x, y) \le n\, D_\infty(x, y), \\ D_\infty(x, y) &\le D_2(x, y) \le \sqrt{n}\, D_\infty(x, y), \end{aligned} \tag{B.6.4}$$

válidas para todos $x, y \in M$, mostram que as métricas D_1, D_2 e D_∞ são duas a duas Lipschitz-equivalentes.

B.6.1. Exemplo. Se para todo $i = 1, \dots, n$, o espaço métrico M_i é $\mathbb{R}$ e d_i é a métrica Euclideana então $M = \prod_{i=1}^n M_i = \mathbb{R}^n$ e as métricas D_1, D_2, D_∞ coincidem respectivamente com as métricas d_1, d e d_∞ definidas no Exemplo B.1.2.

B.6.2. Exercício. Sejam (M_1, d_1), ..., (M_n, d_n) espaços métricos e seja $D : M \times M \to \mathbb{R}$ a métrica produto do tipo 1 associada às métricas d_i, $i = 1, \dots, n$, onde $M = \prod_{i=1}^n M_i$. Para cada $i = 1, \dots, n$, seja N_i um subconjunto de M_i e seja $N = \prod_{i=1}^n N_i$. Mostre que $D|_{N \times N}$ é a métrica produto do tipo 1 associada às métricas $d_i|_{N_i \times N_i}$, $i = 1, \dots, n$. Repita o exercício substituindo todas as ocorrências de "métrica produto tipo 1" por "métrica produto tipo 2" e também por "métrica produto tipo ∞".

B.6.3. Exercício. Sejam (M_{ij}, d_{ij}), $j = 1, \dots, n_i$, $i = 1, \dots, k$, espaços métricos e defina $M_i = \prod_{j=1}^{n_i} M_{ij}$, $i = 1, \dots, k$; seja d_i a métrica produto tipo 1 em M_i associada às métricas d_{ij}, $j = 1, \dots, n_i$, e seja d a métrica produto tipo 1 em $M = \prod_{i=1}^k M_i$ associada às métricas d_i, $i = 1, \dots, k$. Se:

$$M' = \prod_{\substack{1 \le i \le k \\ 1 \le j \le n_i}} M_{ij}$$

é munido da métrica produto tipo 1 associada às métricas d_{ij}, $j = 1, \dots, n_i$, $i = 1, \dots, k$, mostre que a aplicação:

$$\big((x_{11}, \dots, x_{1n_1}), \dots (x_{k1}, \dots, x_{kn_k})\big) \longmapsto (x_{11}, \dots, x_{1n_1}, \dots, x_{k1}, \dots, x_{kn_k})$$

é uma isometria de M sobre M'. Repita o exercício substituindo todas as ocorrências de "métrica produto tipo 1" por "métrica produto tipo 2" e também por "métrica produto tipo ∞".

B.6.4. EXERCÍCIO. Sejam (M_i, d_i), $i = 1, \ldots, n$, espaços métricos e seja $M = \prod_{i=1}^n M_i$ munido de uma das métricas produto associadas às métricas d_i, $i = 1, \ldots, n$. Mostre que as projeções $\pi_i : M \to M_i$ são Lipschitzianas com constante de Lipschitz igual a 1. Conclua que as projeções π_i são uniformemente contínuas.

B.6.5. EXEMPLO. Seja (M, d) um espaço métrico arbitrário e seja D a métrica produto tipo 1 em $M \times M$, isto é:

$$D : (M \times M) \times (M \times M) \ni \big((x, y), (x', y')\big) \longmapsto d(x, x') + d(y, y').$$

Afirmamos que a função $d : M \times M \to \mathbb{R}$ é Lipschitziana com constante de Lipschitz igual a 1; em particular, d é uniformemente contínua. De fato, dados $x, y, x', y' \in M$, temos:

$$d(x, y) \le d(x, x') + d(x', y') + d(y', y),$$

donde:

$$d(x, y) - d(x', y') \le d(x, x') + d(y', y) = D\big((x, y), (x', y')\big).$$

Trocando os papéis de (x, y) e (x', y') obtemos:

$$d(x', y') - d(x, y) \le D\big((x, y), (x', y')\big),$$

donde:

$$\big|d(x, y) - d(x', y')\big| \le D\big((x, y), (x', y')\big).$$

B.6.6. PROPOSIÇÃO. *Sejam (M_1, d_1), ..., (M_n, d_n) espaços métricos e suponha $M = \prod_{i=1}^n M_i$ munido de uma das métricas produto D associadas às métricas d_i, $i = 1, \ldots, n$. Temos que:*

(a) *se A_i é aberto em M_i para todo $i = 1, \ldots, n$, então $A = \prod_{i=1}^n A_i$ é aberto em M;*
(b) *se F_i é fechado em M_i para todo $i = 1, \ldots, n$, então $F = \prod_{i=1}^n F_i$ é fechado em M;*
(c) *se U é aberto em M então para todo $x \in U$ podemos encontrar conjuntos abertos $A_i \subset M_i$, $i = 1, \ldots, n$, tais que $x \in \prod_{i=1}^n A_i \subset U$.*

DEMONSTRAÇÃO. Os itens (a) e (b) seguem da continuidade das projeções $\pi_i : M \to M_i$ (Exercício B.6.4), das igualdades $A = \bigcap_{i=1}^n \pi_i^{-1}(A_i)$, $F = \bigcap_{i=1}^n \pi_i^{-1}(F_i)$ e da Proposição B.2.7. Passemos à prova do item (c); como as métricas produto são todas equivalentes, podemos supor sem perda de generalidade que D é a métrica produto tipo ∞. Dado $x \in U$, como U é aberto em M, existe $r > 0$ tal que $\mathrm{B}(x, r) \subset U$. Mas é fácil ver que:

$$\mathrm{B}(x, r) = \prod_{i=1}^n \mathrm{B}(x_i, r),$$

donde basta tomar $A_i = \mathrm{B}(x_i, r)$, $i = 1, \ldots, n$. $\square$

A proposição a seguir resume todas as propriedades importantes das métricas produto.

B.6.7. PROPOSIÇÃO. *Sejam (M,d), (N_1,d_1), ..., (N_k,d_k), espaços métricos; suponha $N = \prod_{i=1}^k N_i$ munido de algumas das métricas produto D associadas às métricas d_i, $i = 1,\ldots,k$, e denote por $\pi_i : N \to N_i$ a i-ésima projeção. Temos que:*

(a) *uma função $f : M \to N$ é contínua num ponto $x \in M$ se e somente se $\pi_i \circ f$ é contínua no ponto x para todo $i = 1,\ldots,k$;*
(b) *uma função $f : M \to N$ é uniformemente contínua se e somente se $\pi_i \circ f$ é uniformemente contínua para todo $i = 1,\ldots,k$;*
(c) *uma função $f : M \to N$ é Lipschitziana se e somente se $\pi_i \circ f$ é Lipschitziana para todo $i = 1,\ldots,k$;*
(d) *dados um subconjunto A de M, um ponto de acumulação $a \in M$ de A, uma função $f : A \to N$ e um ponto $L \in N$ então L é o limite de f no ponto a se e somente se $\pi_i(L)$ é o limite de $\pi_i \circ f$ no ponto a para todo $i = 1,\ldots,k$;*
(e) *se $(x_n)_{n\geq 0}$ é uma seqüência em N e se $a \in N$ então $(x_n)_{n\geq 0}$ converge para a se e somente se $\big(\pi_i(x_n)\big)_{n\geq 0}$ converge para $\pi_i(a)$ para todo $i = 1,\ldots,k$.*

DEMONSTRAÇÃO. Como todas as métricas produto são Lipschitz-equivalentes, podemos supor sem perda de generalidade que D é a métrica produto tipo ∞. Nesse caso, para todos $p,q \in N$ e para todo $\varepsilon > 0$ temos:

$$D(p,q) < \varepsilon \iff \text{para todo } i = 1,\ldots,k,\ d_i\big(\pi_i(p),\pi_i(q)\big) < \varepsilon.$$

Os itens (a), (b), (d) e (e) seguem quase que diretamente de instâncias adequadas da equivalência acima: usamos $p = f(x)$, $q = f(y)$ na demonstração dos itens (a) e (b), $p = L$, $q = f(x)$ na demonstração do item (d) e $p = x_n$, $q = a$ na demonstração do item (e). Passemos então à demonstração do item (c): se f é Lipschitziana então cada $\pi_i \circ f$ é Lipschitziana, já que π_i é Lipschitziana. Reciprocamente, se cada $\pi_i \circ f$ é Lipschitziana, podemos escolher uma constante de Lipschitz comum $c \geq 0$ para todas as funções $\pi_i \circ f$ (tome o máximo de constantes de Lipschitz para cada uma das funções) e daí:

$$D\big(f(x),f(y)\big) = \max_{i=1,\ldots,n} d_i\big((\pi_i \circ f)(x),(\pi_i \circ f)(y)\big) \leq c\,d(x,y),$$

para todos $x,y \in M$, donde c é uma constante de Lipschitz para f. □

B.7. Conexidade

B.7.1. DEFINIÇÃO. Uma *cisão* de um espaço métrico (M,d) é um par (A,B) de subconjuntos abertos e disjuntos de M tais que $M = A \cup B$. Uma cisão (A,B) é dita *trivial* se $A = \emptyset$ ou $B = \emptyset$. O espaço métrico (M,d) é dito *conexo* se admite apenas a cisão trivial. Um espaço métrico que não é conexo é dito *desconexo*. Um subconjunto X de um espaço métrico (M,d)

é dito *conexo* (resp., *desconexo*) se o espaço métrico $(X, d|_{X\times X})$ é conexo (resp., desconexo).

Evidentemente, cisão e conexidade são conceitos topológicos.

B.7.2. EXERCÍCIO. Seja M um espaço métrico. Mostre que (A, B) é uma cisão de M se e somente se $M = A \cup B$, $A \cap B = \emptyset$ e A, B são fechados.

B.7.3. EXERCÍCIO. Mostre que um espaço métrico M é conexo se e somente se os únicos subconjuntos de M que são ao mesmo tempo abertos e fechados são o conjunto vazio e M.

B.7.4. PROPOSIÇÃO. *Sejam M, N espaços métricos. Se existe uma função contínua sobrejetora $f : M \to N$ e M é conexo então N também é conexo.*

DEMONSTRAÇÃO. Se (A, B) é uma cisão de N então $\big(f^{-1}(A), f^{-1}(B)\big)$ é uma cisão de M, donde $f^{-1}(A) = \emptyset$ ou $f^{-1}(B) = \emptyset$; como f é sobrejetora, segue que $A = \emptyset$ ou $B = \emptyset$. $\square$

B.7.5. COROLÁRIO. *Sejam M, N espaços métricos e $f : M \to N$ uma função contínua. Se X é um subconjunto conexo de M então $f(X)$ é um subconjunto conexo de N.*

DEMONSTRAÇÃO. Basta observar que $f|_X : X \to f(X)$ é uma função contínua e sobrejetora. $\square$

B.7.6. PROPOSIÇÃO. *Sejam M um espaço métrico e X um subconjunto denso de M. Se X é conexo então M é conexo.*

DEMONSTRAÇÃO. Se (A, B) é uma cisão de M então $(A \cap X, B \cap X)$ é uma cisão de X, donde $A \cap X = \emptyset$ ou $B \cap X = \emptyset$. Como X é denso, segue que $A = \emptyset$ ou $B = \emptyset$. $\square$

B.7.7. COROLÁRIO. *Sejam M um espaço métrico, X um subconjunto conexo de M e Y um subconjunto de M. Se $X \subset Y \subset \overline{X}$ então Y é conexo.*

DEMONSTRAÇÃO. Basta ver que X é denso no espaço métrico Y. $\square$

B.7.8. PROPOSIÇÃO. *Seja M um espaço métrico e $(X_i)_{i\in I}$ uma família de subconjuntos conexos de M. Se a interseção $\bigcap_{i\in I} X_i$ é não vazia então a união $\bigcup_{i\in I} X_i$ é conexa.*

DEMONSTRAÇÃO. Escreva $X = \bigcup_{i\in I} X_i$ e seja (A, B) um cisão de X. Escolha qualquer $p \in \bigcap_{i\in I} X_i$. Como $X = A \cup B$, temos $p \in A$ ou $p \in B$; para fixar as idéias, digamos que $p \in A$. Para todo $i \in I$, temos que $(A \cap X_i, B \cap X_i)$ é uma cisão de X_i, donde $B \cap X_i = \emptyset$ (note que $A \cap X_i$ não pode ser vazio, pois $p \in A \cap X_i$). Daí $B = \bigcup_{i\in I}(B \cap X_i) = \emptyset$. $\square$

B.7.9. PROPOSIÇÃO. *Dado um subconjunto I de $\mathbb{R}$, as seguintes condições são equivalentes:*

(a) *I é conexo;*

(b) *para todos $a, b \in I$, $c \in \mathbb{R}$, se $a < c < b$ então $c \in I$.*

Demonstração. Suponha que I é conexo e sejam $a, b, c \in \mathbb{R}$ tais que $a < c < b$ e $a, b \in I$. Se fosse $c \notin I$, teríamos que $\big( ]-\infty, c[\cap I,]c, +\infty[\cap I \big)$ seria uma cisão não trivial de I, o que não é possível; daí $c \in I$. Suponha agora que a condição (b) vale e suponha por absurdo que I seja desconexo. Seja (A, B) uma cisão não trivial de I e sejam $a \in A$, $b \in B$. Temos $a < b$ ou $b < a$; para fixar as idéias, vamos assumir que $a < b$. Considere o conjunto:

$$S = \big\{ x \in [a, +\infty[\, : [a, x] \subset A \big\}.$$

Temos que $a \in S$ e S é limitado superiormente por b, donde S possui um supremo $c \in \mathbb{R}$; de $a \le c \le b$ e da condição (b) vemos que $c \in I$. Como $S \subset A$ e c é aderente a S, temos que c é aderente a A e, como A é fechado em I, temos que $c \in A$. Vemos então que $c < b$, já que $b \in B$. Mas como A é aberto em I, existe $\varepsilon > 0$ com $[c - \varepsilon, c + \varepsilon] \cap I \subset A$; podemos escolher esse ε de modo que $c + \varepsilon \le b$. Daí $a \le c < c + \varepsilon \le b$ e a condição (b) implica que $[c, c + \varepsilon] \subset I$; logo $[c, c + \varepsilon] \subset [c - \varepsilon, c + \varepsilon] \cap I \subset A$. Mas é fácil ver que $[a, c[\subset A$, donde $[a, c + \varepsilon] \subset A$ e $c + \varepsilon \in S$, contradizendo o fato que c é o supremo de S. □

B.7.10. Corolário. $\mathbb{R}$ *é conexo.* □

B.7.11. Exercício. Dizemos que um subconjunto I de $\mathbb{R}$ é um *intervalo* se satisfaz pelo menos uma das condições abaixo:

- $I = \emptyset$ ou $I = \{a\}$, com $a \in \mathbb{R}$;
- $I = [a, b]$ ou $I =]a, b]$ ou $I = [a, b[$ ou $I =]a, b[$, com $a, b \in \mathbb{R}$, $a < b$;
- $I =]-\infty, a]$ ou $I =]-\infty, a[$ ou $I = [a, +\infty[$ ou $I =]a, +\infty[$, com $a \in \mathbb{R}$;
- $I =]-\infty, +\infty[= \mathbb{R}$.

Mostre que um subconjunto de $\mathbb{R}$ é conexo se e somente se é um intervalo (sugestão: mostre que um subconjunto I de $\mathbb{R}$ é um intervalo se e somente se satisfaz a condição (b) no enunciado da Proposição B.7.9; para isso, considere o ínfimo e o supremo de I).

B.7.12. Proposição. *Se M, N são espaços métricos conexos e se $M \times N$ é munido de uma das métricas produto então $M \times N$ é conexo. Reciprocamente, se M e N são não vazios e $M \times N$ é conexo então M e N são conexos.*

Demonstração. Se M e N são não vazios e $M \times N$ é conexo então a conexidade de M e de N segue da Proposição B.7.4 observando que as projeções do produto $M \times N$ são contínuas e sobrejetoras. Suponha agora que M e N são conexos e mostremos que $M \times N$ é conexo. Se M ou N é vazio então $M \times N$ é vazio e portanto conexo; suponha que M e N são não vazios. Dados $x \in M$, $y \in N$, denote por $i_x : N \to M \times N$, $j_y : M \to M \times N$ as aplicações definidas por $i_x(q) = (x, q)$, $j_y(p) = (p, y)$, para todos $p \in M$, $q \in N$. Segue da Proposição B.6.7 que as aplicações i_x e j_y são contínuas e portanto o Corolário B.7.5 nos dá que $i_x(N)$ e $j_y(M)$ são subconjuntos conexos de $M \times N$. Como $(x, y) \in i_x(N) \cap j_y(M)$ vemos que a união $C_{xy} = i_x(N) \cup j_y(M)$ é conexa (Proposição B.7.8). Mas:

$$M \times N = \bigcup_{y \in N} j_y(M) \subset \bigcup_{y \in N} C_{xy} \subset M \times N,$$

donde $M \times N = \bigcup_{y \in N} C_{xy}$; mas:

$$\emptyset \neq i_x(N) \subset \bigcap_{y \in N} C_{xy},$$

donde a Proposição B.7.8 nos dá que $M \times N$ é conexo. □

B.7.13. Corolário. *Se $M_1, \ldots, M_n$ são espaços métricos conexos e $M = \prod_{i=1}^n M_i$ é munido de uma das métricas produto então M é conexo. Reciprocamente, se M é conexo e cada M_i é não vazio então cada M_i é conexo.*

Demonstração. Segue facilmente da Proposição B.7.12 por indução em n. □

B.7.14. Corolário. *$\mathbb{R}^n$ é conexo.*

Demonstração. Segue dos Corolário B.7.10 e B.7.13. □

B.7.15. Definição. Um espaço métrico (M, d) é dito *conexo por caminhos* (ou *conexo por arcos*) se para todos $x, y \in M$ existe uma função contínua $\gamma : [0, 1] \to M$ tal que $\gamma(0) = x$ e $\gamma(1) = y$. Um subconjunto X de M é dito *conexo por caminhos* se o espaço métrico $(X, d|_{X \times X})$ for conexo por caminhos.

Evidentemente, conexidade por caminhos é um conceito topológico.

Note que um subconjunto X de M é conexo por caminhos se e somente se para todos $x, y \in X$ existe uma função contínua $\gamma : [0, 1] \to M$ tal que $\gamma(0) = x$, $\gamma(1) = y$ e $\gamma\big([0, 1]\big) \subset X$.

B.7.16. Proposição. *Todo espaço métrico conexo por caminhos é conexo.*

Demonstração. Seja M um espaço métrico conexo por caminhos. Se M é vazio, então M é conexo. Se M é não vazio, fixe $x \in M$ e denote por $\mathcal{C}$ o conjunto de todas as funções contínuas $\gamma : [0, 1] \to M$ tais que $\gamma(0) = x$. A hipótese de que M é conexo por caminhos implica que:

$$M = \bigcup_{\gamma \in \mathcal{C}} \gamma\big([0, 1]\big).$$

Como cada γ é contínua e $[0, 1]$ é conexo (Proposição B.7.9), segue que $\gamma\big([0, 1]\big)$ é conexo, para todo $\gamma \in \mathcal{C}$ (Corolário B.7.5). Mas como:

$$x \in \bigcap_{\gamma \in \mathcal{C}} \gamma\big([0, 1]\big),$$

segue da Proposição B.7.8 que M é conexo. □

B.8. Completude

B.8.1. Definição. Uma seqüência $(x_n)_{n\geq 0}$ num espaço métrico M é dita *de Cauchy* se para todo $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $d(x_n, x_m) < \varepsilon$, para todos $n, m \in \mathbb{N}$ com $n, m \geq n_0$.

B.8.2. Exercício. Mostre que uma seqüência $(x_n)_{n\geq 0}$ num espaço métrico M é de Cauchy se e somente se para todo $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que o conjunto:

$$\{x_n : n \geq n_0\}$$

tem diâmetro menor do que ε. Conclua que toda seqüência de Cauchy é limitada.

B.8.3. Exercício. Sejam (M, d) um espaço métrico e N um subconjunto de M. Uma seqüência $(x_n)_{n\geq 0}$ em N é de Cauchy no espaço métrico $(N, d|_{N\times N})$ se e somente se ela é de Cauchy em (M, d).

B.8.4. Proposição. *Toda seqüência convergente é de Cauchy.*

Demonstração. Sejam M um espaço métrico e $(x_n)_{n\geq 0}$ uma seqüência em M que converge para $a \in M$. Para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $d(x_n, a) < \frac{\varepsilon}{2}$, para todo $n \geq n_0$. Daí, se $n, m \geq n_0$ temos $d(x_n, a) < \frac{\varepsilon}{2}$ e $d(x_m, a) < \frac{\varepsilon}{2}$, donde a desigualdade triangular nos dá $d(x_n, x_m) < \varepsilon$. $\square$

B.8.5. Exercício. Mostre que uma subseqüência de uma seqüência de Cauchy é ainda uma seqüência de Cauchy.

B.8.6. Proposição. *Sejam M um espaço métrico e $(x_n)_{n\geq 0}$ uma seqüência de Cauchy em M. Se $(x_n)_{n\geq 0}$ possui uma subseqüência $(x_{n_k})_{k\geq 0}$ que converge para um certo $a \in M$ então $(x_n)_{n\geq 0}$ converge para a.*

Demonstração. Seja dado $\varepsilon > 0$ e seja $m_0 \in \mathbb{N}$ tal que $d(x_n, x_m) < \frac{\varepsilon}{2}$, para todos $n, m \in \mathbb{N}$ com $n, m \geq m_0$. Fixe $n \geq m_0$ e vamos mostrar que $d(x_n, a) < \varepsilon$. Como $x_{n_k} \to a$, existe k_0 tal que $d(x_{n_k}, a) < \frac{\varepsilon}{2}$, para todo $k \geq k_0$. Tome $k \geq k_0$ tal que $n_k \geq m_0$. Temos então que $d(x_{n_k}, a) < \frac{\varepsilon}{2}$ e que $d(x_{n_k}, x_n) < \frac{\varepsilon}{2}$, donde a desigualdade triangular nos dá $d(x_n, a) < \varepsilon$. $\square$

B.8.7. Proposição. *Sejam dados espaços métricos (M, d), (N, d'). Se $f : M \to N$ é uma função uniformemente contínua e se $(x_n)_{n\geq 0}$ é uma seqüência de Cauchy em M então $\big(f(x_n)\big)_{n\geq 0}$ é uma seqüência de Cauchy em N.*

Demonstração. Dado $\varepsilon > 0$, existe $\delta > 0$ tal que $d(p, q) < \delta$ implica $d'\big(f(p), f(q)\big) < \varepsilon$, para todos $p, q \in M$; existe também $n_0 \in \mathbb{N}$ tal que $d(x_n, x_m) < \delta$, para todos $n, m \geq n_0$. Daí $d'\big(f(x_n), f(x_m)\big) < \varepsilon$, para todos $n, m \geq n_0$. $\square$

B.8.8. Corolário. *Se d, d' são métricas uniformemente equivalentes em M então uma seqüência $(x_n)_{n\geq 0}$ em M é de Cauchy em (M, d) se e somente se for de Cauchy em (M, d').* $\square$

B.8.9. Exercício. Seja $M =]0, +\infty[$ munido da restrição da métrica Euclideana. Mostre que a função $f : M \to M$ definida por $f(x) = \frac{1}{x}$, para todo $x \in M$ é contínua, que a seqüência $x_n = \frac{1}{n}$, $n = 1, 2, \ldots$, é de Cauchy em M mas que a seqüência $\big(f(x_n)\big)_{n\geq 1}$ não é de Cauchy em M.

B.8.10. Exercício. Sejam (M_1, d_1), ..., (M_n, d_n) espaços métricos e assuma que $M = \prod_{i=1}^{n} M_i$ está munido de uma das métricas produto. Mostre que uma seqüência $(x_k)_{k\geq 0}$ em M é de Cauchy se e somente se para todo $i = 1, \ldots, n$, a seqüência $\big(\pi_i(x_k)\big)_{k\geq 0}$ é de Cauchy em M_i, onde $\pi_i : M \to M_i$ denota a i-ésima projeção.

B.8.11. Definição. Um espaço métrico (M, d) é dito *completo* se toda seqüência de Cauchy em M é convergente. Um subconjunto N de M é dito *completo* se o espaço métrico $(N, d|_{N\times N})$ é completo.

B.8.12. Proposição. *Sejam M um espaço métrico e N um subconjunto de M. Se N é completo então N é um subconjunto fechado de M. Reciprocamente, se M é completo e N é fechado em M então N é completo.*

Demonstração. Suponha que N seja completo. Para ver que N é fechado em M, usaremos o Corolário B.4.19. Seja $(x_n)_{n\geq 0}$ uma seqüência em N que converge em M para $a \in M$; então $(x_n)_{n\geq 0}$ é de Cauchy em M e portanto também é de Cauchy em N. Como N é completo, $(x_n)_{n\geq 0}$ converge em N para um certo $a' \in N$. Mas isso implica que $(x_n)_{n\geq 0}$ converge em M para a', donde $a = a'$ e $a \in N$. Vemos então que N é fechado. Suponha agora que N é fechado e que M seja completo. Seja $(x_n)_{n\geq 0}$ uma seqüência de Cauchy em N; temos que $(x_n)_{n\geq 0}$ é de Cauchy em M e portanto converge em M para um certo $a \in M$. Como N é fechado, temos que $a \in N$ e portanto $(x_n)_{n\geq 0}$ é convergente em N. Isso completa a demonstração de que N é completo. □

B.8.13. Proposição. $\mathbb{R}$ *é completo.*

Demonstração. Seja $(x_n)_{n\geq 0}$ uma seqüência de Cauchy em $\mathbb{R}$. Como toda seqüência de Cauchy é limitada, existe $c \geq 0$ tal que $|x_n| \leq c$, para todo $n \in \mathbb{N}$. Para cada $k \in \mathbb{N}$, seja:

$$y_k = \inf \big\{x_n : n \geq k\big\};$$

temos que esse ínfimo realmente existe e $|y_k| \leq c$, para todo $k \in \mathbb{N}$. Logo existe também o supremo:

$$z = \sup \big\{y_k : k \in \mathbb{N}\big\}.$$

Note que a seqüência $(y_k)_{k\geq 0}$ é crescente. Vamos mostrar que $(x_n)_{n\geq 0}$ converge para z. Seja dado $\varepsilon > 0$ e seja $n_0 \in \mathbb{N}$ tal que:

$$d(x_n, x_m) = |x_n - x_m| < \varepsilon,$$

para todos $n, m \in \mathbb{N}$ com $n, m \geq n_0$. Fixe $k \geq n_0$. Temos $x_n > x_k - \varepsilon$, para todo $n \geq k$ e portanto:

$$z \geq y_k = \inf \big\{x_n : n \geq k\big\} \geq x_k - \varepsilon.$$

Por outro lado, para todo $n \geq n_0$, temos $y_n \leq x_n < x_k + \varepsilon$ e portanto:

$$z = \sup\big\{y_n : n \in \mathbb{N}\big\} = \sup\big\{y_n : n \geq n_0\big\} \leq x_k + \varepsilon.$$

Logo $d(x_k, z) = |x_k - z| \leq \varepsilon$, para todo $k \geq n_0$. □

B.8.14. PROPOSIÇÃO. *Sejam dados espaços métricos* (M, d), (N, d'). *Se existe um homeomorfismo uniformemente contínuo* $f : M \to N$ *e se* N *é completo então* M *é completo.*

DEMONSTRAÇÃO. Seja $(x_n)_{n\geq 0}$ uma seqüência de Cauchy em M. Como f é uniformemente contínuo, a Proposição B.8.7 nos diz que $\big(f(x_n)\big)_{n\geq 0}$ é uma seqüência de Cauchy em N. Como N é completo, essa seqüência de Cauchy converge para um certo $a \in N$; mas aí a continuidade de f^{-1} implica que $(x_n)_{n\geq 0}$ converge para $f^{-1}(a)$. □

B.8.15. COROLÁRIO. *Se* d, d' *são métricas uniformemente equivalentes em* M *e se* (M, d) *é completo então* (M, d') *é completo.* □

B.8.16. EXERCÍCIO. Sejam $M = \,]0, 1]$, $N = [1, +\infty[$ munidos da restrição da métrica Euclideana. Mostre que:

(a) a aplicação $f : M \ni x \mapsto \frac{1}{x} \in N$ é um homeomorfismo;
(b) conclua que a métrica d_f (Exercício B.1.55) é equivalente à métrica de M e que completude (assim como seqüência de Cauchy) não é um conceito topológico (isto é, um conjunto pode ser completo quando munido de uma métrica, mas não ser completo quando munido de outra equivalente).

B.8.17. PROPOSIÇÃO. *Sejam* (M_1, d_1), $\ldots$, (M_n, d_n) *espaços métricos e assuma que* $M = \prod_{i=1}^{n} M_i$ *está munido de uma das métricas produto. Se cada* M_i *é completo então* M *é completo. Reciprocamente, se* M *é completo e se cada* M_i *é não vazio então cada* M_i *é completo.*

DEMONSTRAÇÃO. Assuma que cada M_i é completo e seja $(x_k)_{k\geq 0}$ uma seqüência de Cauchy em M. Se $\pi_i : M \to M_i$ denota a i-ésima projeção então, pelo resultado do Exercício B.8.10, a seqüência $\big(\pi_i(x_k)\big)_{k\geq 0}$ é de Cauchy e portanto converge para um ponto $a_i \in M_i$. Segue então do item (e) da Proposição B.6.7 que a seqüência $(x_k)_{k\geq 0}$ converge para $(a_1, \ldots, a_n) \in M$. Reciprocamente, suponha que M é completo e que cada M_i seja não vazio. Escolha $p_i \in M_i$ para $i = 1, \ldots, k$ e fixe $j \in \{1, \ldots, k\}$. A aplicação:

$$M_j \ni a \longmapsto (p_1, \ldots, p_{j-1}, a, p_{j+1}, \ldots, p_n)$$

é um homeomorfismo uniformemente contínuo sobre o subconjunto fechado:

$$\{p_1\} \times \cdots \times \{p_{j-1}\} \times M_j \times \{p_{j+1}\} \times \cdots \times \{p_n\}$$

de M (Proposição B.6.6). Segue então das Proposições B.8.12 e B.8.14 que M_j é completo. □

B.8.18. COROLÁRIO. $\mathbb{R}^n$ *é completo.*

DEMONSTRAÇÃO. Segue das Proposições B.8.13 e B.8.17. □

B.8.19. Proposição. *Seja M um espaço métrico. Temos que M é completo se e somente se vale a seguinte condição: para toda seqüência decrescente*[6] $(F_n)_{n\geq 0}$ *de fechados não vazios com* $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$*, a interseção* $\bigcap_{n=0}^{\infty} F_n$ *é não vazia.*

Demonstração. Suponha que M é completo e seja $(F_n)_{n\geq 0}$ uma seqüência decrescente de fechados não vazios com $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$. Para cada $n \in \mathbb{N}$, escolha $x_n \in F_n$. Afirmamos que a seqüência $(x_n)_{n\geq 0}$ é de Cauchy. De fato, dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $\operatorname{diam}(F_{n_0}) < \varepsilon$; daí, se $n, m \geq n_0$ então $x_n \in F_n \subset F_{n_0}$ e $x_m \in F_m \subset F_{n_0}$, donde:

$$d(x_n, x_m) \leq \operatorname{diam}(F_{n_0}) < \varepsilon.$$

Seja $a = \lim_{n\to\infty} x_n$. Para todo $n \in \mathbb{N}$, a subseqüência $(x_{n+k})_{k\geq 0}$ de $(x_n)_{n\geq 0}$ converge para a e $x_{n+k} \in F_n$, para todo $k \in \mathbb{N}$; como F_n é fechado, segue que $a \in F_n$ e portanto $a \in \bigcap_{n=0}^{\infty} F_n \neq \emptyset$. Suponha agora que para toda seqüência decrescente $(F_n)_{n\geq 0}$ de fechados não vazios com $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$ a interseção $\bigcap_{n=0}^{\infty} F_n$ é não vazia e vamos mostrar que M é completo. Seja $(x_n)_{n\geq 0}$ uma seqüência de Cauchy em M. Para todo $n \geq 0$, sejam:

$$A_n = \big\{ x_k : k \geq n \big\}$$

e $F_n = \overline{A_n}$. Como, para todo $n \in \mathbb{N}$, $A_n \supset A_{n+1}$, temos $F_n \supset F_{n+1}$; obviamente, F_n é não vazio, para todo $n \in \mathbb{N}$. Segue facilmente do resultado do Exercício B.8.2 que $\lim_{n\to\infty} \operatorname{diam}(A_n) = 0$ e da Proposição B.1.42 segue então que $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$. Pelas nossas hipóteses, existe $a \in \bigcap_{n=0}^{\infty} F_n$. Vamos mostrar que $x_n \to a$. Dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $\operatorname{diam}(F_{n_0}) < \varepsilon$; como $a \in F_{n_0}$, temos $d(x_n, a) \leq \operatorname{diam}(F_{n_0}) < \varepsilon$, para todo $n \geq n_0$. Isso completa a demonstração. □

B.9. Compacidade

Seja M um conjunto. Uma *cobertura* de M é uma família $(U_i)_{i\in I}$ de conjuntos tal que $M = \bigcup_{i\in I} U_i$. Uma *subcobertura* de uma cobertura $(U_i)_{i\in I}$ é uma subfamília $(U_i)_{i\in J}$, $J \subset I$, de $(U_i)_{i\in I}$ tal que $M = \bigcup_{i\in J} U_i$. Se (M, d) é um espaço métrico então uma família $(U_i)_{i\in I}$ de subconjuntos de M é dita *aberta* se U_i é aberto em M, para todo $i \in I$.

B.9.1. Definição. Um espaço métrico (M, d) é dito *compacto* se toda cobertura aberta de M admite uma subcobertura finita. Um subconjunto K de M é dito *compacto* se o espaço métrico $(K, d|_{K\times K})$ é compacto.

Evidentemente compacidade é um conceito topológico.

Dado um subconjunto K de M, então uma família $(U_i)_{i\in I}$ de subconjuntos de M com $K \subset \bigcup_{i\in I} U_i$ é dita uma *cobertura de K por subconjuntos de M*.

---

[6] Isto é, $F_n \supset F_{n+1}$, para todo $n \in \mathbb{N}$.

B.9.2. Proposição. *Seja M um espaço métrico e seja K um subconjunto de M. Temos que K é compacto se e somente se toda cobertura $(U_i)_{i\in I}$ de K por subconjuntos abertos de M admite uma subcobertura finita (isto é, existe $J \subset I$ finito com $K \subset \bigcup_{i\in J} U_i$).*

Demonstração. Suponha que K é compacto e seja $(U_i)_{i\in I}$ uma cobertura de K por subconjuntos abertos de M. Daí $(K \cap U_i)_{i\in I}$ é uma cobertura aberta de K e portanto existe $J \subset I$ finito tal que $K = \bigcup_{i\in J}(K \cap U_i)$; logo $K \subset \bigcup_{i\in J} U_i$. Reciprocamente, suponha que toda cobertura de K por subconjuntos abertos de M admite uma subcobertura finita e vamos mostrar que K é compacto. Seja $(V_i)_{i\in I}$ uma cobertura aberta de K; para cada $i \in I$, existe U_i aberto em M com $V_i = K \cap U_i$. Daí $K \subset \bigcup_{i\in I} U_i$ e pelas nossas hipóteses existe $J \subset I$ finito com $K \subset \bigcup_{i\in J} U_i$. Logo $K = \bigcup_{i\in J}(K \cap U_i) = \bigcup_{i\in J} V_i$. □

B.9.3. Definição. Um espaço métrico (M, d) é dito *seqüêncialmente compacto* se toda seqüência em M admite uma subseqüência convergente. Um subconjunto K de M é dito *seqüêncialmente compacto* se o espaço métrico $(K, d|_{K\times K})$ é seqüêncialmente compacto.

Evidentemente, um espaço métrico M é seqüêncialmente compacto se e somente se toda seqüência em M possui algum valor de aderência (Definição B.4.14).

Veremos daqui a pouco que compacidade e compacidade seqüêncial são propriedades equivalentes para um espaço métrico[7].

B.9.4. Proposição. *Seja M um espaço métrico e seja A um subconjunto de M. Temos que $\overline{A}$ é seqüêncialmente compacto se e somente se toda seqüência em A admite uma subseqüência que é convergente em M.*

Demonstração. Se $\overline{A}$ é seqüêncialmente compacto então toda seqüência em A é uma seqüência em $\overline{A}$ e portanto admite uma subseqüência que é convergente em $\overline{A}$; essa subseqüência também é convergente em M. Reciprocamente, suponha que toda seqüência em A admite uma subseqüência que é convergente em M e seja $(x_n)_{n\geq 1}$ uma seqüência em $\overline{A}$. Vamos mostrar que $(x_n)_{n\geq 1}$ possui uma subseqüência convergente em $\overline{A}$. Para todo $n \geq 1$, a bola aberta $\mathrm{B}\left(x_n, \frac{1}{n}\right)$ intercepta A e portanto existe $y_n \in A$ com $d(x_n, y_n) < \frac{1}{n}$. Pelas nossas hipóteses, a seqüência $(y_n)_{n\geq 1}$ admite uma subseqüência $(y_{n_k})_{k\geq 0}$ que converge para algum $a \in M$; como $(y_{n_k})_{k\geq 0}$ é uma seqüência em A, temos que $a \in \overline{A}$. Vamos mostrar que a subseqüência $(x_{n_k})_{k\geq 0}$ de $(x_n)_{n\geq 1}$ converge para a. Dado $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$ tal que $d(y_{n_k}, a) < \frac{\varepsilon}{2}$, para todo $k \geq k_0$; podemos escolher k_0 de modo que $n_{k_0} > \frac{2}{\varepsilon}$. Daí, para todo $k \geq k_0$, temos:

$$d(x_{n_k}, a) \leq d(x_{n_k}, y_{n_k}) + d(y_{n_k}, a) < \frac{1}{n_k} + \frac{\varepsilon}{2} < \varepsilon. \qquad \square$$

[7] Essas propriedades não são equivalentes para espaços topológicos, mas nós não estudaremos espaços topológicos aqui.

B.9.5. DEFINIÇÃO. Um espaço métrico (M, d) é dito *totalmente limitado* se para todo $\varepsilon > 0$ existe uma cobertura finita de M por conjuntos de diâmetro menor do que ε. Um subconjunto A de M é dito *totalmente limitado* se o espaço métrico $(A, d|_{A\times A})$ é totalmente limitado.

Evidentemente todo espaço métrico totalmente limitado é limitado.

B.9.6. EXERCÍCIO. Mostre que todo subconjunto de um espaço métrico totalmente limitado é totalmente limitado (em particular, um subconjunto de um subconjunto totalmente limitado é totalmente limitado).

B.9.7. EXERCÍCIO. Se M é um espaço métrico e A é um subconjunto de M, mostre que A é totalmente limitado se e somente se $\overline{A}$ é totalmente limitado (sugestão: use a Proposição B.1.34 e a Proposição B.1.42).

B.9.8. PROPOSIÇÃO. *Seja M um espaço métrico. São equivalentes as seguintes condições:*

(a) *M é compacto;*
(b) *M é seqüêncialmente compacto;*
(c) *M é completo e totalmente limitado.*

DEMONSTRAÇÃO. Assuma que M é compacto e vamos mostrar que uma seqüência arbitrária $(x_n)_{n\geq 0}$ em M possui algum valor de aderência. Suponha por absurdo que tal valor de aderência não exista; daí, para todo $a \in M$, como a não é um valor de aderência para $(x_n)_{n\geq 0}$, a Proposição B.4.17 nos diz que existe um aberto U_a contendo a tal que o conjunto:

$$\big\{n \in \mathbb{N} : x_n \in U_a\big\}$$

é finito. Temos então que $(U_a)_{a\in M}$ é uma cobertura aberta de M, da qual podemos extrair uma subcobertura finita, i.e., existe um subconjunto finito F de M tal que $M = \bigcup_{a\in F} U_a$. Mas daí obtemos que:

$$\mathbb{N} = \bigcup_{a\in F} \big\{n \in \mathbb{N} : x_n \in U_a\big\}$$

donde chegamos ao absurdo de que $\mathbb{N}$ é um conjunto finito. Logo $(x_n)_{n\geq 0}$ possui algum valor de aderência e nós provamos que M é seqüêncialmente compacto. Suponha agora que M é seqüêncialmente compacto e vamos demonstrar que M é completo e totalmente limitado. Em primeiro lugar, segue diretamente da Proposição B.8.6 que M é completo. Suponha por absurdo que M não seja totalmente limitado; daí existe $\varepsilon > 0$ tal que M não admite uma cobertura finita por subconjuntos de diâmetro menor do que ε. Vamos construir recursivamente uma seqüência $(x_n)_{n\geq 0}$ em M da seguinte forma: tome qualquer ponto $x_0 \in M$ (M não é vazio, pois M não é totalmente limitado). Supondo $x_0, \ldots, x_n \in M$ definidos, observamos que:

$$\bigcup_{i=0}^{n} \mathrm{B}\big(x_i, \tfrac{\varepsilon}{3}\big) \neq M,$$

já que $\operatorname{diam}\big[\mathrm{B}\big(x_i, \frac{\varepsilon}{3}\big)\big] \leq \frac{2\varepsilon}{3} < \varepsilon$. Escolha qualquer $x_{n+1} \in M$ tal que x_{n+1} não pertence a $\bigcup_{i=0}^{n} \mathrm{B}\big(x_i, \frac{\varepsilon}{3}\big)$. Obtemos assim uma seqüência $(x_n)_{n\geq 0}$ em

M com $d(x_n, x_m) \geq \frac{\varepsilon}{3}$, para todos $n, m \in \mathbb{N}$ com $n \neq m$. Essa seqüência não pode ter subseqüência convergente, pois ela não tem nenhuma subseqüência de Cauchy. Isso contradiz a hipótese de que M é seqüêncialmente compacto e mostra que M é totalmente limitado. Suponhamos, por fim, que M seja completo e totalmente limitado e mostremos que M é compacto. Seja $(U_i)_{i\in I}$ uma cobertura aberta de M. Suponha por absurdo que $(U_i)_{i\in I}$ não admite uma subcobertura finita. Nós construiremos recursivamente uma seqüência decrescente $(F_n)_{n\geq 1}$ de conjuntos fechados não vazios tal que para todo $n \geq 1$, $\operatorname{diam}(F_n) < \frac{1}{n}$ e F_n não pode ser coberto por um número finito de conjuntos U_i. Em primeiro lugar, como M é totalmente limitado, M pode ser coberto por um número finito de conjuntos de diâmetro menor do que 1; ao menos um desses conjuntos, chamemos ele de A_1, não pode ser coberto por um número finito de conjuntos U_i. Denote por F_1 o fecho de A_1. Temos $\operatorname{diam}(F_1) = \operatorname{diam}(A_1) < 1$ e F_1 não pode ser coberto por um número finito de conjuntos U_i. Suponha que tenha sido definido o conjunto fechado F_n que não pode ser coberto por um número finito de conjuntos U_i e vamos definir F_{n+1}. Como F_n é totalmente limitado, F_n é igual à união de um número finito de conjuntos de diâmetro menor do que $\frac{1}{n+1}$; ao menos um desses conjuntos, chamemos ele de A_{n+1}, não pode ser coberto por um número finito de conjuntos U_i. Seja F_{n+1} o fecho de A_{n+1}. Temos então que $F_{n+1} \subset F_n$, $\operatorname{diam}(F_{n+1}) < \frac{1}{n+1}$ e F_{n+1} não pode ser coberto por um número finito de conjuntos U_i. Obtivemos então uma seqüência decrescente $(F_n)_{n\geq 1}$ de conjuntos fechados tal que, par todo $n \geq 1$, $\operatorname{diam}(F_n) < \frac{1}{n}$ e F_n não pode ser coberto por um número finito de conjuntos U_i; em particular, $F_n \neq \emptyset$. Como M é completo e $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$, a Proposição B.8.19 nos dá um elemento $a \in \bigcap_{n=1}^{\infty} F_n$. Como $M = \bigcup_{i\in I} U_i$, existe $i_0 \in I$ com $a \in U_{i_0}$ e como U_{i_0} é aberto existe $r > 0$ tal que $\mathrm{B}(a, r) \subset U_{i_0}$. Seja $n \geq 1$ tal que $\operatorname{diam}(F_n) < r$. Como $a \in F_n$, temos $F_n \subset \mathrm{B}(a, r) \subset U_{i_0}$, contradizendo o fato de que F_n não pode ser coberto por um número finito de conjuntos U_i. Isso prova que M é compacto e completa a demonstração. □

B.9.9. Corolário. *Todo espaço métrico compacto é limitado e todo subconjunto compacto de um espaço métrico é fechado.*

Demonstração. Segue do fato que todo espaço métrico totalmente limitado é limitado e do fato que todo subconjunto completo de um espaço métrico é fechado (Proposição B.8.12). □

B.10. Bases de abertos e separabilidade

B.10.1. Definição. Seja M um espaço métrico. Uma *base de abertos* para M é uma coleção $\mathfrak{B}$ de subconjuntos abertos para M tal que para todo conjunto aberto U existe um subconjunto $\mathfrak{S} \subset \mathfrak{B}$ tal que $U = \bigcup_{B\in\mathfrak{S}} B$.

B.10.2. Exercício. Sejam (M, d) um espaço métrico, N um subconjunto de M e $\mathfrak{B}$ uma base de abertos para M. Mostre que $\big\{B \cap N : B \in \mathfrak{B}\big\}$ é uma base de abertos para o espaço métrico $(N, d|_{N\times N})$.

B.10.3. PROPOSIÇÃO. *Seja M um espaço métrico. Uma coleção $\mathfrak{B}$ de conjuntos abertos é uma base de abertos para M se e somente se para todo conjunto aberto U e para todo $x \in U$ existe $B \in \mathfrak{B}$ com $x \in B$ e $B \subset U$.*

B.10.4. DEFINIÇÃO. Um espaço métrico M é dito *separável* se admite um subconjunto enumerável denso.

B.10.5. DEFINIÇÃO. Dizemos que um espaço métrico M é um *espaço de Lindelöf* se toda cobertura aberta de M admite uma subcobertura enumerável.

B.10.6. PROPOSIÇÃO. *Seja M um espaço métrico. São equivalentes:*

(a) *M admite uma base de abertos enumerável;*
(b) *M é separável;*
(c) *M é um espaço de Lindelöf.*

DEMONSTRAÇÃO. □

B.11. Espaços vetoriais normados

No que segue, todos os espaços vetoriais considerados são *reais*.

B.11.1. DEFINIÇÃO. Seja E um espaço vetorial. Uma *norma* em E é uma aplicação $E \ni x \mapsto \|x\| \in \mathbb{R}$ que satisfaz as seguintes condições:

(a) $\|x\| \geq 0$, para todo $x \in E$;
(b) para todo $x \in E$, se $\|x\| = 0$ então $x = 0$;
(c) $\|\lambda x\| = |\lambda|\,\|x\|$, para todos $\lambda \in \mathbb{R}$, $x \in E$;
(d) (desigualdade triangular) $\|x+y\| \leq \|x\|+\|y\|$, para todos $x, y \in E$.

Um *espaço vetorial normado* é um espaço vetorial munido de uma norma (ou, mais precisamente, um par ordenado formado por um espaço vetorial e uma norma nesse espaço vetorial).

Quando não há possibilidade de confusão, muitas vezes denotaremos mais de uma norma (às vezes normas em espaços diferentes) pelo mesmo símbolo $\|\cdot\|$.

B.11.2. EXERCÍCIO. Se $\|\cdot\|$ é uma norma num espaço vetorial E, mostre que:

$$d : E \times E \ni (x, y) \longmapsto \|x - y\| \in \mathbb{R}$$

é uma métrica em E.

A métrica definida no Exercício B.11.2 é chamada a *métrica induzida* pela norma $\|\cdot\|$. A menos de menção explícita em contrário, assuminos que todo espaço vetorial normado está munido da métrica induzida pela sua norma.

B.11.3. EXEMPLO. A *norma Euclideana* em $\mathbb{R}^n$ é definida por:

$$\|x\| = \Big(\sum_{i=1}^{n} x_i^2\Big)^{\frac{1}{2}},$$

para todo $x \in \mathbb{R}^n$ (a desigualdade triangular para $\|\cdot\|$ segue do resultado do Exercício A.1 aplicado ao produto interno canônico de $\mathbb{R}^n$). Também são normas em $\mathbb{R}^n$:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|, \quad \|x\|_\infty = \max\big\{|x_i|; i = 1, \ldots, n\big\}, \qquad x \in \mathbb{R}^n.$$

A norma $\|\cdot\|_1$ é também conhecida como *norma da soma* e a norma $\|\cdot\|_\infty$ é também conhecida como *norma do máximo*. A norma Euclideana $\|\cdot\|$ é às vezes também denotada por $\|\cdot\|_2$. É claro que as métricas induzidas por $\|\cdot\|_1$, $\|\cdot\|$ e $\|\cdot\|_\infty$ são, respectivamente, as métricas d_1, d e d_∞ (veja Exemplo B.1.2). No caso particular em que $n = 1$, todas essas normas coincidem com o valor absoluto.

B.11.4. Exemplo. Sejam $(E_i, \|\cdot\|_i)$, $i = 1, \ldots, n$, espaços vetoriais normados e considere o produto cartesiano (ou soma direta externa) $E = \prod_{i=1}^n E_i$. São normas em E:

$$E \ni x \longmapsto \sum_{i=1}^{n} \|x_i\|_i \in \mathbb{R}, \tag{B.11.1}$$

$$E \ni x \longmapsto \Big(\sum_{i=1}^{n} \|x_i\|^2\Big)^{\frac{1}{2}} \in \mathbb{R}, \tag{B.11.2}$$

$$E \ni x \longmapsto \max\big\{\|x_i\|_i : i = 1, \ldots, n\big\} \in \mathbb{R}. \tag{B.11.3}$$

Se cada E_i é munido da métrica d_i induzida pela norma $\|\cdot\|_i$ então as métricas induzidas em E pelas normas (B.11.1), (B.11.2) e (B.11.3) são, respectivamente, as métricas D_1, D_2 e D_∞ definidas em (B.6.1), (B.6.2) e (B.6.3). Obviamente, se cada E_i é $\mathbb{R}$ e $\|\cdot\|_i$ é o valor absoluto então $E = \mathbb{R}^n$ e as normas (B.11.1), (B.11.2) e (B.11.3) são iguais, respectivamente, às normas da soma, Euclideana e do máximo em $\mathbb{R}^n$ (veja Exemplo B.11.3 acima).

B.11.5. Definição. Duas normas $\|\cdot\|$, $\|\cdot\|'$ num espaço vetorial E são ditas *equivalentes* (resp., Lipschitz-equivalentes) se as métricas induzidas pelas mesmas são equivalentes (resp., Lipschitz-equivalentes).

B.11.6. Exercício. Mostre que duas normas $\|\cdot\|_1$, $\|\cdot\|_2$ num espaço vetorial E são Lipschitz-equivalentes se e somente se existem números reais $k, k' > 0$ tais que:

$$k'\|x\|_1 \le \|x\|_2 \le k\|x\|_1, \tag{B.11.4}$$

para todo $x \in E$ (sugestão: para provar (B.11.4) a partir de (B.5.1) faça $y = 0$ em (B.5.1) e para provar (B.5.1) a partir de (B.11.4) troque x por $x - y$ em (B.11.4)).

B.11.7. Exemplo. As normas em $\mathbb{R}^n$ definidas no Exemplo B.11.3 são duas a duas Lipschitz-equivalentes (veja Exemplo B.5.11), o mesmo sendo verdade para as normas definidas no Exemplo B.11.4 (veja (B.6.4)).

Seguirá da próxima proposição que duas normas num espaço vetorial são equivalentes se e somente se são Lipschitz-equivalentes.

B.11.8. PROPOSIÇÃO. *Sejam E, F espaços vetoriais normados e seja $T : E \to F$ uma aplicação linear. As seguintes condições são equivalentes:*

(a) *T é contínua;*
(b) *T é contínua na origem;*
(c) *existe $k \geq 0$ tal que $\|T(x)\| \leq k$, para todo $x \in E$ com $\|x\| \leq 1$;*
(d) *existe $k \geq 0$ tal que $\|T(x)\| \leq k\|x\|$, para todo $x \in E$;*
(e) *T é Lipschitziana.*

DEMONSTRAÇÃO. É óbvio que (a) implica (b). Assumindo (b) então, dado $\varepsilon = 1$, existe $\delta > 0$ tal que $\|T(x)\| \leq 1$, para todo $x \in E$ com $\|x\| \leq \delta$. Daí, se $\|x\| \leq 1$, temos $\|\delta x\| \leq \delta$, donde:

$$\|T(x)\| = \frac{1}{\delta}\,\|T(\delta x)\| \leq \frac{1}{\delta},$$

o que prova (c). Assuindo (c) então, para todo $x \in E$ com $x \neq 0$ temos:

$$\frac{1}{\|x\|}\,\|T(x)\| = \Big\|T\Big(\frac{x}{\|x\|}\Big)\Big\| \leq k,$$

donde $\|T(x)\| \leq k\|x\|$; essa desigualdade é óbvia para $x = 0$, e isso prova (d). Assumindo (d), temos, para todos $x, y \in E$:

$$\|T(x) - T(y)\| = \|T(x - y)\| \leq k\|x - y\|,$$

donde T é Lipschitziana. Finalmente, é óbvio que (e) implica (a). □

B.11.9. COROLÁRIO. *Duas normas $\|\cdot\|$, $\|\cdot\|'$ num espaço vetorial E são equivalentes se e somente se são Lipschitz-equivalentes.*

DEMONSTRAÇÃO. A aplicação identidade de $(E, \|\cdot\|)$ para $(E, \|\cdot\|')$ (assim como sua inversa) é linear e portanto é Lipschitziana se e somente se for contínua. □

B.11.10. OBSERVAÇÃO. É fácil ver que uma constante $k \geq 0$ satisfaz a condição (c) no enunciado a Proposição B.11.8, se e somente se satisfaz a condição (d) no enunciado dessa proposição, se e somente se for uma constante de Lipschitz para T. A menor constante que satisfaz essas condições (se existir) será o supremo de $\|T(x)\|$ com $\|x\| \leq 1$; nós definimos então, se $T : E \to F$ é uma aplicação linear contínua, a *norma* de T por:

$$\|T\| = \sup_{\|x\| \leq 1} \|T(x)\|. \tag{B.11.5}$$

Temos então:

$$\|T(x)\| \leq \|T\|\,\|x\|,$$

para todo $x \in E$.

B.11.11. Exercício. Dados espaços vetoriais normados E, F, mostre que o conjunto de todas as aplicações lineares contínuas de E para F é um subespaço do espaço vetorial $\mathrm{Lin}(E, F)$ de todas as aplicações lineares de E para F e que (B.11.5) define uma norma nesse subespaço. Se E, F, G são espaços vetoriais normados e $T : E \to F$, $S : F \to G$ são aplicações lineares contínuas, mostre que:

$$\|S \circ T\| \le \|S\| \, \|T\|.$$

B.11.12. Exemplo. Seja E um espaço vetorial. A soma de E:

$$E \times E \ni (x, y) \longmapsto x + y \in E \tag{B.11.6}$$

é uma aplicação linear (não bilinear!). Escolha uma norma $\|\cdot\|$ em E. Usando a norma $(x, y) \mapsto \|x\| + \|y\|$ em $E \times E$, então a desigualdade triangular para $\|\cdot\|$ nos mostra que (B.11.5) é uma aplicação linear contínua, já que é Lipschitziana com constante de Lipschitz 1.

B.11.13. Exercício. Sejam M um espaço métrico e E um espaço vetorial normado. Dadas funções $f : M \to E$, $g : M \to E$, podemos definir $f + g : M \to E$ fazendo $(f + g)(x) = f(x) + g(x)$, para todo $x \in M$. Mostre que se f, g são contínuas num ponto $x \in M$ então $f + g$ também é contínua no ponto x. Mostre também que se f, g são uniformemente contínuas (resp., Lipschitzianas) então também $f + g$ é uniformemente contínua (resp., Lipschitziana).

A Proposição B.11.8 possui uma generalização natural para aplicações multilineares.

B.11.14. Proposição. *Sejam E_1, ..., E_n, F espaços vetoriais normados e seja $B : E_1 \times \cdots \times E_n \to F$ uma aplicação multilinear. Se usamos em $E_1 \times \cdots \times E_n$ uma das normas equivalentes definidas no Exemplo B.11.4 então as seguintes condições são equivalentes:*

(a) *B é contínua;*
(b) *B é contínua na origem;*
(c) *existe $k \ge 0$ tal que:*

$$\|B(x_1, \ldots, x_n)\| \le k,$$

para todos $x_1 \in E_1$, ..., $x_n \in E_n$, com $\|x_i\| \le 1$, $i = 1, \ldots, n$;

(d) *existe $k \ge 0$ tal que:*

$$\|B(x_1, \ldots, x_n)\| \le k\|x_1\| \cdots \|x_n\|, \tag{B.11.7}$$

para todos $x_1 \in E_1$, ..., $x_n \in E_n$.

Demonstração. É óbvio que (a) implica (b). Assumindo (b) então, dado $\varepsilon = 1$, existe $\delta > 0$ tal que:

$$\max\big\{\|x_1\|, \ldots, \|x_n\|\big\} \le \delta$$

implica $\|B(x_1, \ldots, x_n)\| \le 1$, para todos $x_1 \in E_1$, ..., $x_n \in E_n$. Daí, se $\|x_i\| \le 1$, temos $\|\delta x_i\| \le \delta$ e portanto:

$$\|B(x_1, \ldots, x_n)\| = \frac{1}{\delta^n} \|B(\delta x_1, \ldots, \delta x_n)\| \le \frac{1}{\delta^n},$$

para todos $x_1 \in E_1, \ldots, x_n \in E_n$ com $\|x_i\| \le 1$, $i = 1, \ldots, n$; isso prova (c). Assumindo (c), temos:

$$\frac{1}{\|x_1\| \cdots \|x_n\|} \|B(x_1, \ldots, x_n)\| = \Big\| B\Big(\frac{x_1}{\|x_1\|}, \ldots, \frac{x_n}{\|x_n\|}\Big) \Big\| \le k,$$

donde (B.11.7) vale se todos os x_i são não nulos. Mas (B.11.7) é óbvia se algum $x_i = 0$, o que prova (d). Finalmente, vamos assumir (d) e provar que B é contínua. Sejam $x_1 \in E_1, \ldots, x_n \in E_n$ fixados e seja dado $\varepsilon > 0$; devemos determinar $\delta > 0$ tal que se $h_1 \in E_1, \ldots, h_n \in E_n$ satisfazem $\|h_i\| < \delta$, $i = 1, \ldots, n$, então:

$$\|B(x_1 + h_1, \ldots, x_n + h_n) - B(x_1, \ldots, x_n)\| < \varepsilon.$$

Usando a multilinearidade de B, vê-se que $B(x_1 + h_1, \ldots, x_n + h_n)$ é igual a uma soma de 2^n termos, cada um deles da forma $B(\alpha_1, \ldots, \alpha_n)$, onde cada α_i é x_i ou h_i; o termo em que $\alpha_1 = x_1, \ldots, \alpha_n = x_n$ cancela-se quando fazemos a diferença $B(x_1 + h_1, \ldots, x_n + h_n) - B(x_1, \ldots, x_n)$. Seja $M > 0$ tal que $\|x_i\| \le M$, $i = 1, \ldots, n$; temos:

$$\|B(\alpha_1, \ldots, \alpha_n)\| \le k\delta^r M^{n-r},$$

onde r é o número de índices i tais que α_i é h_i. Se $\delta \le M$ e $r \ge 1$, teremos $k\delta^r M^{n-r} \le k\delta M^{n-1}$ e portanto:

$$\|B(x_1 + h_1, \ldots, x_n + h_n) - B(x_1, \ldots, x_n)\| \le k(2^n - 1)\delta M^{n-1}.$$

A demonstração é concluída tomando $\delta > 0$ pequeno o suficiente de modo que $\delta \le M$ e $k(2^n - 1)\delta M^{n-1} < \varepsilon$. □

B.11.15. Observação. Uma constante $k \ge 0$ satisfaz a condição (c) no enunciado da Proposição B.11.14 se e somente se satisfaz a condição (d) no enunciado dessa proposição (mas, diferentemente do caso de aplicações lineares, a constante k não é em geral uma constante de Lipschitz!). A menor constante k que satisfaz essas condições é o supremo de $\|B(x_1, \ldots, x_n)\|$, com $\|x_i\| \le 1$, $i = 1, \ldots, n$. Nós definimos então, se $B : E_1 \times \cdots \times E_n \to F$ é uma aplicação multilinear contínua, a *norma* de B fazendo:

(B.11.8) $$\|B\| = \sup_{\substack{\|x_i\| \le 1 \\ i = 1, \ldots, n}} \|B(x_1, \ldots, x_n)\|.$$

Obviamente:

$$\|B(x_1, \ldots, x_n)\| \le \|B\| \, \|x_1\| \cdots \|x_n\|,$$

para todos $x_1 \in E_1, \ldots, x_n \in E_n$.

B.11.16. Exercício. Dados espaços vetoriais normados $E_1, \ldots, E_n$, F, mostre que o conjunto de todas as aplicações multilineares contínuas de $E_1 \times \cdots \times E_n$ para F é um subespaço do espaço vetorial $\mathrm{Lin}(E_1, \ldots, E_n; F)$ de todas as aplicações multilineares de $E_1 \times \cdots \times E_n$ para F e que (B.11.8) define uma norma nesse subespaço.

B.11.17. EXEMPLO. Seja E um espaço vetorial. A multiplicação por escalar de E:

$$\mathbb{R} \times E \ni (\lambda, x) \longmapsto \lambda x \in E \tag{B.11.9}$$

é uma aplicação bilinear. Escolha uma norma $\|\cdot\|$ em E. Se usamos no produto cartesiano $\mathbb{R} \times E$ uma das normas usuais (veja Exemplo B.11.4) então segue da Proposição B.11.14 e da identidade $\|\lambda x\| = |\lambda| \, \|x\|$ que a aplicação bilinear (B.11.9) é contínua.

B.11.18. EXERCÍCIO. Sejam M um espaço métrico e E um espaço vetorial normado. Dadas funções $f : M \to \mathbb{R}$, $g : M \to E$, podemos definir uma aplicação $fg : M \to E$ fazendo $(fg)(x) = f(x)g(x)$, para todo $x \in M$. Mostre que se f, g são contínuas num ponto $x \in M$ então também fg é contínua no ponto x.

B.11.1. Normas em espaços vetoriais de dimensão finita. Nossa meta agora é demonstrar as duas seguintes proposições:

B.11.19. PROPOSIÇÃO. *Se E é um espaço vetorial de dimensão finita então quaisquer duas normas em E são equivalentes (e, na verdade, pelo Corolário B.11.9, são também Lipschitz-equivalentes).*

B.11.20. PROPOSIÇÃO. *Se E, F são espaços vetoriais normados e E tem dimensão finita então toda aplicação linear $T : E \to F$ é contínua. Mais geralmente, se $E_1, \ldots, E_n$ são espaços vetoriais normados de dimensão finita então toda aplicação multilinear $B : E_1 \times \cdots \times E_n \to F$ é contínua.*

Em primeiro lugar, observamos que a Proposição B.11.19 segue diretamente da Proposição B.11.20:

DEMONSTRAÇÃO DA PROPOSIÇÃO B.11.19. Se $\|\cdot\|$, $\|\cdot\|'$ são normas em E então a aplicação identidade de $(E, \|\cdot\|)$ para $(E, \|\cdot\|')$ é linear e portanto contínua, pela Proposição B.11.20; o mesmo vale, evidentemente, para sua aplicação inversa. $\square$

Para demonstrar a Proposição B.11.20, usamos o seguinte:

B.11.21. LEMA. *Quaisquer duas normas em $\mathbb{R}^n$ são equivalentes.*

DEMONSTRAÇÃO. Por transitividade, basta mostrar que uma norma arbitrária $\|\cdot\|$ em $\mathbb{R}^n$ é equivalente à norma da soma $\|\cdot\|_1$ (Exemplo B.11.3). Observamos em primeiro lugar que para todo $x \in \mathbb{R}^n$, temos:

$$\|x\| = \Big\| \sum_{i=1}^n x_i e_i \Big\| \le \sum_{i=1}^n |x_i| \, \|e_i\| \le k \|x\|_1,$$

onde $(e_i)_{i=1}^n$ denota a base canônica de $\mathbb{R}^n$ e $k = \max\{\|e_1\|, \ldots, \|e_n\|\}$. Daí a aplicação identidade de $(\mathbb{R}^n, \|\cdot\|_1)$ para $(\mathbb{R}^n, \|\cdot\|)$ é Lipschitziana. Para mostrar que sua inversa é Lipschitziana, devemos mostrar que existe $r \ge 0$ tal que $\|x\|_1 \le r$, para todo $x \in \mathbb{R}^n$ com $\|x\| \le 1$. Se tal r não existisse então para todo número natural r haveria um $x_r \in \mathbb{R}^n$ com $\|x_r\| \leq 1$ e $\|x_r\|_1 > r$. Seja:

$$y_r = \frac{x_r}{\|x_r\|_1}, \quad r = 0, 1, 2, \ldots.$$

Como a seqüência $(y_r)_{r\geq 0}$ é limitada com respeito a $\|\cdot\|_1$, existe uma subseqüência $(y_{r_i})_{i\geq 0}$ que converge, com respeito a $\|\cdot\|_1$, para um certo $y \in \mathbb{R}^n$; note que $\|y\|_1 = 1$, já que $\|y_r\|_1 = 1$, para todo r. Como a aplicação identidade de $(\mathbb{R}^n, \|\cdot\|_1)$ para $(\mathbb{R}^n, \|\cdot\|)$ é contínua, segue que $(y_{r_i})_{i\geq 0}$ converge para y também com respeito a $\|\cdot\|$; mas, por outro lado:

$$\|y_r\| = \frac{\|x_r\|}{\|x_r\|_1} < \frac{1}{r} \xrightarrow[r\to\infty]{} 0,$$

donde $(y_{r_i})_{i\geq 0}$ converge para zero com respeito a $\|\cdot\|$. Concluímos então que $y = 0$, contradizendo $\|y\|_1 = 1$. □

B.11.22. COROLÁRIO. *Se E é um espaço vetorial normado então toda aplicação linear $T : \mathbb{R}^n \to E$ é contínua, sendo $\mathbb{R}^n$ munido de uma norma arbitrária.*

DEMONSTRAÇÃO. O Lema B.11.21 nos permite assumir que os espaços $\mathbb{R}^n$ e $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_k} \cong \mathbb{R}^{n_1+\cdots+n_k}$ estão munidos (por exemplo) da norma do máximo (veja Exemplo B.11.3). Usando a norma do máximo em $\mathbb{R}^n$ então as projeções $\pi_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, n$, são contínuas, já que $|\pi_i(x)| \leq \|x\|_\infty$, para todo $x \in \mathbb{R}^n$. Denote por $(e_i)_{i=1}^n$ a base canônica de $\mathbb{R}^n$. A continuidade de T segue da fórmula:

$$T(x) = \sum_{i=1}^{n} \pi_i(x) T(e_i), \quad x \in \mathbb{R}^n,$$

e também da continuidade de somas e produtos de funções contínuas (veja Exercícios B.11.13 e B.11.18). □

DEMONSTRAÇÃO DA PROPOSIÇÃO B.11.20. Começamos por demonstrar a continuidade de uma aplicação linear $T : E \to F$, onde E tem dimensão finita n. Seja $\phi : \mathbb{R}^n \to E$ um isomorfismo linear arbitrário; podemos definir uma norma $\|\cdot\|'$ em $\mathbb{R}^n$ que torna ϕ uma isometria, fazendo:

$$\|x\|' = \|\phi(x)\|, \quad x \in \mathbb{R}^n,$$

onde $\|\cdot\|$ denota a norma de E. Segue do Corolário B.11.22 que $T \circ \phi$ é contínua; como ϕ é um homeomorfismo, concluímos que T é contínua. □

Exercícios para o Apêndice B

Definição e conceitos básicos.

B.1. EXERCÍCIO. Seja $\phi : [0, +\infty[\to [0, +\infty[$ uma função crescente (isto é, $a \leq b$ implica $\phi(a) \leq \phi(b)$) que satisfaz a desigualdade:

$$\phi(a + b) \leq \phi(a) + \phi(b),$$

para todos $a, b \geq 0$. Suponha que, para todo $a \in [0, +\infty[$, $\phi(a) = 0$ se e somente se $a = 0$. Se (M, d) é um espaço métrico, mostre que a função $\phi \circ d$ também é uma métrica em M.

B.2. Exercício. Mostre que as funções:

$$[0, +\infty[\; \ni a \longmapsto \min\{a, 1\} \in [0, +\infty[\,,$$
$$[0, +\infty[\; \ni a \longmapsto \frac{a}{1+a} \in [0, +\infty[\,,$$

satisfazem as condições que aparecem no enunciado do Exercício B.1. Conclua que se (M, d) é um espaço métrico então:

$$M \times M \ni (x, y) \longmapsto \min\big\{d(x, y), 1\big\}, \tag{B.10}$$

$$M \times M \ni (x, y) \longmapsto \frac{d(x, y)}{1 + d(x, y)}, \tag{B.11}$$

também são métricas em M.

Funções contínuas e uniformemente contínuas.

B.3. Exercício. Seja $\phi : [0, +\infty[\to [0, +\infty[$ uma função que satisfaz as condições que aparecem no enunciado do Exercício B.1. Mostre que se ϕ é contínua no ponto 0 então a aplicação identidade de (M, d) para $(M, \phi \circ d)$ é uniformemente contínua. Se existe $a > 0$ tal que a função $\phi|_{[0,a]}$ é injetora e se a função $(\phi|_{[0,a]})^{-1} : \phi\big([0, a]\big) \to [0, a]$ é contínua no ponto 0, mostre que também a aplicação identidade de $(M, \phi \circ d)$ para (M, d) é uniformemente contínua.

Topologia e equivalência entre métricas.

B.4. Exercício. Se (M, d) é um espaço métrico, mostre que as métricas (B.10) e (B.11) são uniformemente equivalentes a d (sugestão: use o resultado do Exercício B.3). Conclua que toda métrica é uniformemente equivalente a uma métrica limitada.

Métricas no produto cartesiano.

B.5. Exercício. Seja $\phi : [0, +\infty[^n \to [0, +\infty[$ uma função satisfazendo as seguintes propriedades:

(a) se $0 \leq a_1 \leq b_1, \ldots, 0 \leq a_n \leq b_n$ então $\phi(a) \leq \phi(b)$;
(b) para todo $a \in [0, +\infty[^n$, temos $\phi(a) = 0$ se e somente se $a = 0$;
(c) para todos $a, b \in [0, +\infty[^n$, temos $\phi(a + b) \leq \phi(a) + \phi(b)$.

Se $(M_1, d_1), \ldots, (M_n, d_n)$ são espaços métricos e se $M = \prod_{i=1}^n M_i$, mostre que:

$$M \times M \ni (x, y) \longmapsto \phi\big(d_1(x_1, y_1), \ldots, d_n(x_n, y_n)\big) \in \mathbb{R}$$

é uma métrica em M. Mostre também que as funções:

$$[0,+\infty[^n \ni a \longmapsto \sum_{i=1}^{n} a_i \in [0,+\infty[\,, \tag{B.12}$$

$$[0,+\infty[^n \ni a \longmapsto \Big(\sum_{i=1}^{n} a_i^2\Big)^{\frac{1}{2}} \in [0,+\infty[\,, \tag{B.13}$$

$$[0,+\infty[^n \ni a \longmapsto \max\{a_1,\ldots,a_n\} \in [0,+\infty[\,, \tag{B.14}$$

satisfazem as condições (a), (b) e (c) acima (sugestão: para provar que (B.13) satisfaz (c) use o resultado do item (b) do Exercício A.1 aplicado ao produto interno canônico de $\mathbb{R}^n$). Conclua que as aplicações D_1, D_2 e D_∞ definidas em (B.6.1), (B.6.2) e (B.6.3) são realmente métricas.

Referências Bibliográficas

[1] D. J. Bernstein, *Calculus for mathematicians*, http://cr.yp.to/papers/calculus.pdf