Coisas que o Luís precisa aprender

Marino Abasolo, R.

École Normale Supérieure

13 de junho de 2011

1 Introdução

Luís. Não é muito normal receber textos desse tipo, eu confesso nunca ter recebido um, mas decidi tomar um tempo de meu tédio para escrever essas poucas linhas sobre um assunto bem interessante que não sei se já viu, imagino que não. Há muito que você precisa aprender e é meu trabalho garantir que você saiba essas coisas o mais rápido possível, você é meu aprendiz de segunda ordem e devo garantir que você deve não apenas ser melhor que eu, mas que meus aprendizes de primeira ordem, trabalho duro, pois eles são bons.

Escrever algo para eles me tomaria muito tempo e trabalho, eles agora lidam com assuntos que, para que eu escreva, precisaria de cuidado, esmero e tempo; com você as coisas são mais fáceis. Decidi escrever essas poucas linhas sobre algumas bases matemáticas de algumas coisas que você já viu, mas que provavelmente nunca parou para pensar, e garanto que sempre vale a pena parar para pensar. Confesso que comecei esse texto leve, mas depois me empolguei e acabei colocando coisas mais avançadas, mas nada que você não consiga entender se não parar para pensar um pouco, nem precisa ser muito. Nessa matéria, acostumar-se com o que se está fazendo é o mais importante. No entanto, ela representa talvez uma primeira apresentação de "matemática de verdade", um verdadeiro desafio mental honesto e belo. Divirta-se.

2 Motivação

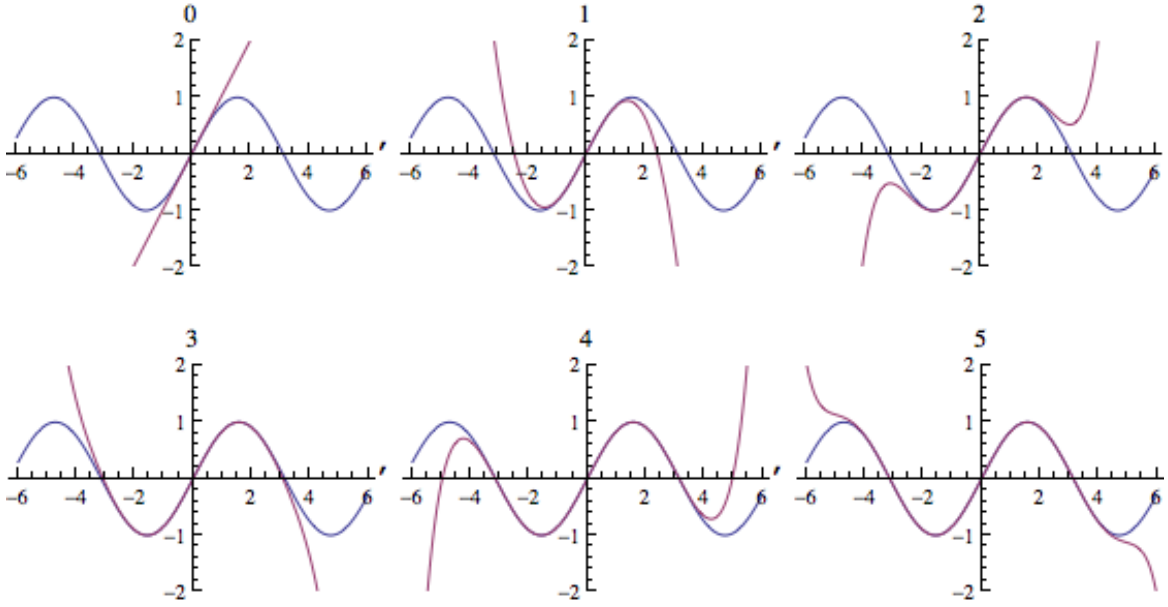
Não duvido de que você saiba o que é uma derivada e para que serve, não me arrisco a ensiná-lo isso. Suponho também que saiba o que é uma integral e saiba fazer algumas, se não sabe, deveria saber. Mas quero falar um pouco de matemática de verdade, e essa matemática não é feita de números, mas de ideias. Vou começar falando sobre a série de Taylor.

Se ainda não ouviu falar dela, certamente já viu aquela famosa série para o cálculo de $\sin(x)$:

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \cdots$$

Se não conhecia essa expressão, é um passatempo divertido. Calcule-a para valores conhecidos do seno, compare os valores, a velocidade de aproximação dela é impressionante. Tão

impressionante que vale a pena ver como ela se aproxima conforme eu somo termos à série. Até coloquei uma figura para você entender do que estou falando.



Parece seguro dizer que, quanto mais termos eu somo, mais esse polinômio fica próximo da função seno. Mas, então vem a pergunta, o que exatamente quer dizer uma função ser próxima da outra?

Isso é um problema complicado. Para se divertir, você calculava para valores conhecidos e checava que eles se aproximavam muito, mas você não pode se dar ao luxo de dizer que uma função é próxima da outra se todos os seus valores são próximos, ora, você não pode calcular todos! O gráfico dá uma boa indicação dessa proximidade, você pode até pensar em calcular a diferença de cada ponto e depois somar. Ora, isso não me parece inteligente, pois há muitos pontos, muitos mesmo.

Outro problema. Suponha você que eu tenha uma matriz bem especial, chamada matriz de rotação. Ela tem uma forma parecida com: $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. Quando θ é bem pequeno, essa matriz tem que estar bem próxima da identidade, ou seja, rodar quase nada tem que estar bem próximo de não rodar absolutamente nada. Além do já visto problema de proximidade de funções, tenho a pergunta: o que é uma matriz estar próxima da outra?

Um problema parecido com esse, mas na física. Na dinâmica das rotações, teremos que uma quantidade conhecida como momento angular é proporcional à velocidade angular, mas não necessariamente diretamente proporcional, pois a o corpo pode ter uma forma tal que rotações em determinadas direções "contém" mais que rotações em outras direções. Tanto a velocidade angular quanto o momento angular são vetores, mas a relação deles se dá através de uma matriz, chamada tensor de inércia I, da forma:

$$\vec{L} = I\vec{\omega}$$

Nesse caso, como avaliar qual corpo possui "mais" momento de inércia? Nesse caso, teríamos que avaliar quem tem mais I, mas como definir o tamanho de uma matriz?

Todos esses problemas estão relacionados, todos envolvem o conceito matemático de distância e tamanho, algo que você não deve ter visto com muito carinho e rigor, então me pus a escrever este breve texto para esclarecer essa noção para você, ou para deixar tudo mais arcano. Tudo o que eu citar aqui possui um número tão vasto de aplicações na física e na matemática que me foi difícil escolher exemplos, essa é a base de diversas áreas matemáticas e é absolutamente fundamental para qualquer coisa em física. Claro, se você estivesse em engenharia, poderia passar direto, eu ia falar que a matriz é grande e a função é próxima. Mas qualquer ciência de verdade não pode se contentar com isso.

3 Distâncias

3.1 Definições

Vamos começar falando de distâncias. Você deve ter percebido que é de meu interesse deixar as coisas o mais geral possível, porque na minha lista de exemplos eu poderia pegar funções, vetores, matrizes, operadores lineares autoadjuntos, cofunctores, qualquer coisa. Por isso, vou dizer que os elementos cuja distância quero medir estão em um conjunto X.

Já dá para imaginar que uma distância entre dois elementos de X é uma função que leva $(x,y) \in X$ aos reais, nada muito surpreendente até agora. Não quero distâncias imaginárias, nem distâncias negativas, então $d : X \times X \to \mathbb{R}^+$.

E quais propriedades essa função distância deveria ter, para que possa ser chamada de distância? Pensemos nas mais razoáveis. É de bom tom que a distância de um ponto a si mesmo seja nula. Ora, não é exigir muito. Quero também que a distância seja nula apenas entre pontos que são o mesmo, ou seja, caras diferentes devem possuir uma distância entre si. Seria também interessante que a distância entre x e y seja a mesma que a distância entre y e x.

Essa propriedade parece evidente no nosso sentido de distância convencional, usando uma fita métrica para medir, mas é fácil perceber que se tentarmos definir como distância o tempo que levamos para chegar a algum lugar, a distância entre o centro de São Paulo e a periferia será bem diferente de fazer o caminho inverso. Como terceira propriedade, podemos pensar que a distância entre x e y deve ter um valor menor que a soma das distâncias entre x e z com a distância entre z e y, intuitivamente, essa distância seria a medida do "menor caminho" entre dois pontos. Se um outro caminho, passando por z, fosse menor que o menor caminho, eu teria problemas. Temos então a lista, ou, aliás, a definição:

Definição 1 (Distância)**.** *Seja X um conjunto. Dizemos que $d : X \to \mathbb{R}$ é uma distância sobre X se ela satisfaz:*

- $d(x,y) = 0 \Leftrightarrow x = y$.
- $d(x,y) = d(y,x)$
- $d(x,y) \leq d(x,z) + d(z,y)$

Essa última propriedade é conhecida como desigualdade triangular, pela forte analogia que ela possui com a condição de existência de um triângulo. Para que o triângulo exista, as distâncias entre seus vértices deve obedecer a essa terceira propriedade. Como estou em um bom humor invejável, até coloquei uma figura de triângulo para você entender do que estou falando.

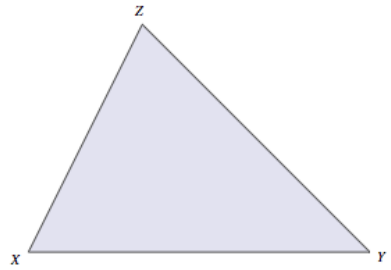


Figura 1: Um triângulo, oras.

E isso basta para definir uma distância. É bem pouco, se você parar para pensar, muita coisa satisfaz minha propriedade de distância. Na verdade, isso implica que existem diversos jeitos de definir distância entre objetos que eu já conheço. Não é algo ruim, eu agora posso desenvolver uma teoria toda sem se importar com o que é minha distância e tirar conclusões que valem para qualquer distância possível.

Vale a pena citar uma definição aqui. Toda vez que tratamos de um conjunto, vamos munir ele de uma distância. Assim, é natural dar um nome a conjuntos munidos com distâncias, e é o que faremos:

Definição 2 (Espaço métrico)**.** *O conjunto X munido de uma distância d é dito um espaço métrico, denotado (X,d), ou apenas X quando a distância está subentendida.*

Outra noção importante é a ideia de um conjunto limitado. Nos reais você já imagina o que seja isso, são conjuntos cujos elementos não ficam indefinidamente grandes (diferentemete de $\mathbb{N}$, $\mathbb{Z}$ ou $\mathbb{Q}$). No caso dos espaços métricos, não possuímos a noção de um elemento "maior" que o outro, não podemos usar, então, a palavra "grande". Mas podemos usar a palavra "distante". Um conjunto será limitado que a maior distância possível entre dois elementos for finita, isso parece bem razoável.

Definição 3 (Espaço métrico limitado)**.** *Seja X um espaço métrico. Dizemos que X é limitado se a maior distância possível entre dois elementos for um número finito D, chamado diâmetro do conjunto. Em outras palavras, sempre é possível encontrar um valor $M \in \mathbb{R}$ tal que não exista nenhum par de elementos tais que $d(x,y) > M$.*

3.2 Resultados

Você já deve ter visto a definição de limite, mas duvido que tenha visto limites com matrizes, funções ou cofunctores. É para isso que serve essa generalização, eu vou definir limite da maneira mais geral possível.

Definição 4 (Limite de uma função)**.** *Dados dois espaços métricos X e Y munidos de distâncias d_X e d_Y. Seja f uma função que os relaciona: $f : X \to Y$. Dizemos que o limite de f é $L \in Y$ quando $x \in X$ tende a $a \in X$ se:*

$$\forall \varepsilon > 0, \exists \delta > 0 \mid d_X(x,a) < \delta \implies d_Y(f(x),L) < \varepsilon$$

À primeira vista isso é bem assustador, mas lendo com calma fica bem intuitivo. A ideia de limite está associada à ideia de proximidade, queremos dizer que quando x se aproxima de a, então $f(x)$ se aproxima se L. Mas o que a palavra "próximo" quer dizer?

Imagine que você coloca sua mão perto da parede. Em relação ao tamanho do seu corpo, ela pode ser uma distância pequena, em relação a distâncias interatômicas, sua mão está bem longe da parede. Perto é relativo, mas você pode escolher o que é perto para você. Digamos que você diga que perto são 2cm, ou 2mm de distância. Se uma garota fala para você: "não chega perto de mim!", você entende que perto são dois metros, ou vinte. Podendo escolher qual é a distância que define a expressão "estar próximo", queremos escrever que quando x está próximo de a, $f(x)$ fica próximo de L. Para que essa definição seja algo sério, ela precisa ser verdade para qualquer valor de "próximo que eu escolher.

Então eu digo que o limite de $f(x)$ é L se eu posso falar: "eu consigo encontrar um x suficientemente próximo de a tal que $f(x)$ fique próximo de L". Mas quanto é esse próximo? Eu escolho: 1. Então eu quero dizer que sempre posso encontrar um x suficientemente próximo de a de forma a deixar $f(x)$ e L a uma distância menor que 1. Mas eu poderia ter escolhido uma distância de 0,1 para considerar os valores próximos. Então eu ainda consigo encontrar um x ainda mais próximo de a tal que $f(x)$ e L fiquem a uma distância menor que 0,1 dentre si. O limite vai te dizer que qualquer valor de próximo que você pegar, consegue sempre achar um x próximo o suficiente de a para satisfazer a seu critério, por mais exigente que você seja em sua ideia de proximidade.

Assim, vou escrever: não importa o quão rigoroso eu seja com a proximidade, eu sempre encontro um x próximo o suficiente de a tal que $f(x)$ esteja àquela distância "próxima" de L. Isso significa que o limite de f quando x tende a a é L. Mas eu preciso escrever em matemática

isso. Dessa forma:

$$\begin{aligned}\text{Não importa o quão rigoroso eu seja com a proximidade,} &= \forall \varepsilon > 0 \\ \text{eu sempre encontro um } x \text{ próximo o suficiente de } a &= \exists \delta > 0 \mid d_X(x,a) < \delta \\ \text{tal que } f(x) \text{ esteja àquela distância “próxima” de } L. &= \implies d_Y(f(x),L) < \varepsilon\end{aligned}$$

Com essa definição, podemos também nos divertir definindo o limite de uma sequência de elementos. Ora, você bem sabe que a sequência $1/n$ tende a zero conforme n aumenta (em geral, n sempre aumenta e é natural, isso fica subentendido). Mais uma vez, a pergunta: o que é uma sequência tender a um elemento? E mais uma vez usamos o bom senso: dizemos que a_n tende a L quando podemos pensar em uma noção de próximo e encontrar um valor para n tão grande que, a partir dele, a sequência está próxima.

Definição 5 (Limite de uma sequência)**.** *Dado um espaço métrico X munido de uma distância d_X. Seja a_n uma sequência de elementos de X. Dizemos que o limite dessa sequência é L se:*

$$\forall \varepsilon > 0, \exists M \in \mathbb{N} \mid n > M \implies d_X(a_n,L) < \varepsilon$$

Excelente. Agora podemos começar a nos divertir mostrando propriedades desses limites e dessas sequências. A vantagem, mais uma vez, é o quão genéricas nossas hipóteses são. Logo mais, mostro aplicações dessa variedade de possibilidades, ou seja, os mesmos teoremas em contextos completamente diferentes.

Antes, precisamos de um teorema importante. Ele é conhecido como o “teorema do confronto”, *squeeze theorem* na literatura em inglês. Ele lida apenas com funções reais, mas nossas distâncias são valores em $\mathbb{R}$, ele será extremamente útil no que segue:

Teorema 1 (do Confronto)**.** *Dadas três funções reais (que possuem como contradomínio os reais) g,f,h e uma relação de ordem que vale para todos os pontos de seu domínio:*

$$g(x) \leq f(x) \leq h(x), \forall x$$

Suponha que $\lim_{x\to a} g(x) = \lim_{x\to a} h(x) = L$*, então* $\lim_{x\to a} f(x) = L$.

Qual a ideia do teorema? Se uma função está espremida entre duas outras e essas duas vão para o mesmo lugar, a do meio é forçada a ir para o mesmo lugar. Havia um professor russo famoso em minha faculdade por explicar esse teorema como “dois guardas levando o bandido para a prisão”. Tendo ele sido educado na união soviética, tenho medo de pensar na imagem mental que ele tinha desse teorema. Por isso, prefiro que você tenha essa figura como imagem mental, duas funções convergindo para um ponto e forçando a que está entre elas a segui-la. Ora, não é como se ela tivesse muita escolha.

Mas não basta explicar, ou convencer, precisa provar. E é isso que vou fazer. Essa não é uma prova fácil, e fico triste de ter que mandar essa como primeira demonstração, mas esse teorema é uma ferramenta fundamental para a maior parte das outras demonstrações então preciso muito dele. Fazer uma chave de fenda não é fácil, mas uma vez que eu a fiz o resto fica bem mais tranquilo de ser feito.

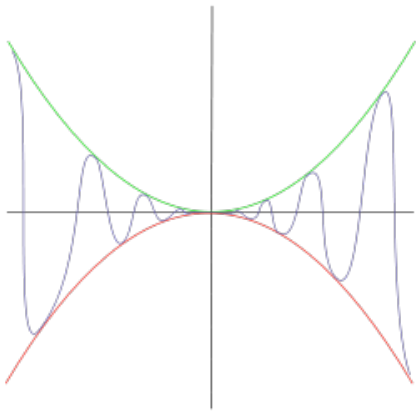


Figura 2: Ilustração do teorema do confronto

E se uma demonstração não parece óbvia para você, ou você acha que "jamais teria pensado nisso", saiba que isso é normal. Você não teria pensado, agora pensaria. Demonstrações matemáticas são, no fundo, um conjunto de truques e maneiras de pensar que você utiliza na ordem e na forma certa. Com o tempo, você aprende mais truques e manobras mais sofisticadas, é um treinamento bem diferente de calcular equações no colegial ou multiplicar matrizes, é algo que exige pensar.

Demonstração. Vou usar uma técnica de demonstração bem útil: provar o teorema para um caso mais "fraco" e depois provar, a partir do caso fraco, o forte. Vou supor que $g(x) \equiv 0$ e $L = 0$, depois mostro que na verdade podia ter escolhido qualquer valor para eles.

Lema 1. *Dadas duas funções reais f,h e uma relação de ordem que vale para todos os pontos de seu domínio:*

$$0 \leq f(x) \leq h(x), \forall x$$

Se $\lim_{x\to a} h(x) = 0$*, então* $\lim_{x\to a} f(x) = 0$.

Demonstração do Lema. Ora, se o limite de $h(x)$ quando x tende a a é zero, então dá pra achar um δ tal que, quando $d(x,a) \leq \delta$, teremos $|h(x)| < \varepsilon$. Note que aqui já posso usar a distância convencional nos reais, que é o módulo da diferença. Vou deixar essa pergunta por aqui para quem sabe depois responder: "a demonstração ainda valeria se eu escolhesse outra distância?". A resposta é sim, mas depois provo isso.

Se $0 \leq f(x) \leq h(x)$ e $h(x) < \varepsilon$, então $f(x) < \varepsilon$ se $d(x,a) \leq \delta$ também, que é a definição do limite para $f(x)$. Então, porque isso é exatamente a definição de limite, $\lim_{x\to a} f(x) = 0$. □

Provado o caso particular, vamos provar que o caso geral se reduz ao particular. Eu tenho a expressão: $g(x) \leq f(x) \leq h(x)$. Eu subtraio $g(x)$ de todo mundo, a desigualdade ainda vale: $0 \leq f(x) - g(x) \leq h(x) - g(x)$. Basta provar que esse é igual ao caso particular que o teorema está provado. Tomemos $h(x) - g(x)$. Sabemos que os limites de h e g são L, então eu consigo

achar $\delta > 0$ tal que:

$$0 \leq |h(x) - g(x)| \overset{1}{=} |h(x) - L + L - g(x)| \overset{2}{\leq} |h(x) - L| + |g(x) - L| \overset{3}{<} \varepsilon + \varepsilon = 2\varepsilon$$

Numerei as passagens para justificar e explicar cada uma delas.

1. Aqui, eu somei zero de um jeito estranho $(L - L)$. Esse é o *oldest trick in the book*, arrisco dizer que provar teoremas é quase a arte de somar zero e multiplicar por um de maneiras pouco óbvias. Se você precisa fazer um cara aparecer, some-o e subtraia-o, jogue o $+$ para um lado e o $-$ para outro, você acaba chegando lá. Foi o que fiz.

2. Eu poderia provar essa desigualdade para você, mas prefiro notar que a função módulo $|x - y|$ é uma distância, ela satisfaz à desigualdade triangular e foi ela que usei. Com distâncias, esse é praticamente o único truque permitido nas regras oficiais, ele é a terceira propriedade das distâncias. Por fim, usei o fato de $d(x,y) = d(y,x)$ para reordenar $|L - g(x)|$ como $|g(x) - L|$.

3. Aqui, usei a definição de limite, mas há uma sutileza. Se g e h possuem L como limite quando x tende a a, isso significa que, dado um $\varepsilon > 0$:

$$\exists \delta_g > 0 \mid d_X(x,a) < \delta_g \implies |g(x) - L| < \varepsilon$$
$$\exists \delta_h > 0 \mid d_X(x,a) < \delta_h \implies |h(x) - L| < \varepsilon$$

   Note que os δ's são diferentes, claro, funções diferentes possuem "exigências" diferentes no nível de proximidade para satisfazer meus critérios de o que é estar perto (meu ε escolhido). Assim, para conseguir usar essa desigualdade eu preciso tomar um $\delta = \min\{\delta_g, \delta_h\}$, o menor dentre esses dois valores de delta. Assim, garanto que tanto $|g(x) - L|$ quanto $|h(x) - L|$ serão menores que o ε. Isso prova que posso escrever aquela desigualdade.

Esses três itens são mais importantes do que parecem. Cada um deles é uma técnica fundamental de demonstração, não só nessa área: somar zero, desigualdade triangular e tomar o menor ou maior valor de um majorante ou minorante para satisfazer a duas condições simultâneas.

Provei que o limite do cara à direita da desigualdade é zero, porque aquela grande desigualdade em que fiz os três passos, tirando essa parte do meio, é exatamente a definição de limite (com 2ε no lugar de ε, o que dá no mesmo, já que posso pegar qualquer ε). O limite do cara da direita é zero, porque ele é sempre zero. Pelo lema, o resultado particular, eu sei que $|f(x) - g(x)| \to 0$ quando $x \to a$. Vamos agora provar que isso implica que o limite de f será L também.

Ora, basta somar zero em $|f(x) - L|$ e repetir alguns dos truques acima:

$$0 \leq |f(x) - L| = |f(x) - g(x) + g(x) - L| \leq |f(x) - g(x)| + |g(x) - L| < 2\varepsilon$$

E provamos que $\lim_{x \to a} f(x) = L$. □

É, deu trabalho, mas vale a pena. O próximo teorema tem a demonstração exatamente igual ao anterior, então não vou fazer, tanto não vou que nem teorema ele será chamado, mas corolário, que é uma consequência do teorema anterior:

Corolário 1 (Teorema do confronto para sequências)**.** *Dadas três sequências reais a_n,b_n,c_n e uma relação de ordem que vale para todos os seus elementos:*

$$a_n \leq b_n \leq c_n, \forall n$$

Se $a_n \to L$ e $c_n \to L$, então $b_n \to L$.

Ele é um subcaso do teorema anterior, uma sequência real nada mais é que uma função real com $\mathbb{N}$ no domínio. Como provamos para qualquer domínio, esse teorema é apenas um corolário do anterior. Ele é extremamente poderoso e as técnicas que introduzi em sua demonstração vão nos ajudar bastante para o próximo teorema:

Teorema 2 (Unicidade do Limite)**.** *O limite de uma função, se existe, é único.*

Demonstração. Suponhamos então que o limite não é único. Suponha que existam dois elementos L e L' tais que, dado $\varepsilon > 0$:

$$\exists \delta_1 > 0 \mid d_X(x,a) < \delta_1 \implies d_Y(f(x),L) < \varepsilon$$
$$\exists \delta_2 > 0 \mid d_X(x,a) < \delta_2 \implies d_Y(f(x),L') < \varepsilon$$

Eu então tomo um $\delta = \min\{\delta_1,\delta_2\}$ para garantir que os dois serão satisfeitos ao mesmo tempo. Isso significa que:

$$0 \leq d_Y(L,L') \leq d_Y(f(x),L) + d_Y(f(x),L') < 2\varepsilon$$

Ora, isso significa que, qualquer que seja o valor de ε que eu tome, por menor que ele seja, eu terei que $d_Y(L,L') < 2\varepsilon$. Vamos provar que isso significa que $L = L'$, o que parece bem intuitivo. Para fazer isso com algum rigor, tomamos uma sequência de ε's! Vamos pegar valores do tipo $\varepsilon_n = 1/n$, uma sequência que tem zero como limite. Teremos uma desigualdade da forma:

$$0 < d_Y(L,L') < 1/n$$

Ora, a sequência da esquerda é a sequência constante (0,0,0...). Não sei se posso elevar ao status de teorema a frase "toda sequência $a_n = a$ constante possui limite e seu limite é a", porque sua prova seria apenas dizer que, para qualquer n, qualquer mesmo, $d(a_n,a) = d(a,a) = 0 < \varepsilon$ para qualquer ε. Então temos que a da esquerda é uma sequência constante com limite 0.

A do meio é também uma sequência constante. resta saber seu limite.

A da direita é uma sequência com limite 0. Ora, **pelo teorema do confronto**, o limite da sequência do meio é zero. Sendo ela constante, todos os seus termos devem coincidir com o limite, logo todos são nulos, $d_Y(L,L') = 0$. Pela primeira propriedades das distâncias, isso implica que $L = L'$. Assim, se existem dois limites, eles são iguais. Logo, o limite é único. □

3.3 Exemplos de distâncias

Vamos agora nos divertir definindo distâncias. Você já conhece uma forma de distância entre números reais, por exemplo, que é a função módulo. Nós a usamos bastante lá em cima, confesso que sem muito rigor na justificativa da razão de a ter escolhido para o teorema do confronto. Não vou entrar nesse nível de detalhes, eu mesmo me perdi nas justificativas dos meandros dessa escolha e alguns colegas meus também, chegamos à conclusão que uma resposta digna exigiria cavar muito fundo nas definições de "maior que" e até na de número real, o que não vem ao caso.

E no caso das matrizes? Um jeito legal de definir distância entre duas matrizes reais é somar a distância de cada termo a cada termo. Sejam M e N duas matrizes com elementos $\{m_{ij}\}$ e $\{n_{ij}\}$, eu defino:

$$d(M,N) = \sum_{i,j} |m_{ij} - n_{ij}|$$

Não faço aqui, mas você pode demonstrar que essa distância de fato satisfaz todas as propriedades de distância que exijo. Para ser mais exato, é mais interessante falar que essa distância *herda* a propriedade de distância da distância $|x - y|$ dos reais. O módulo da diferença é distância, é fácil ver que essa soma de módulos também o será.

No caso do $\mathbb{R}^2$ a situação fica interessante. A distância "canônica", a padrão, entre dois pontos no $\mathbb{R}^2$ é dada por $d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ e você pode se divertir provando que o pitágoras é de fato uma boa distância. Mas eu posso ser criativo, porque minha teoria vale para qualquer distância. Vou usar como distância entre dois pontos no $\mathbb{R}^2$ $x = (x_1,x_2)$ e $y = (y_1,y_2)$ a soma da diferença dos módulos termo a termo: $d_1(x,y) = |x_1 - y_1| + |x_2 - y_2|$, ela também é distância pois herda as propriedades do módulo.

Uma outra distância possível em $\mathbb{R}^2$ é a denotada distância d_∞, ela toma como distância entre x e y o a maior distância entre seus componentes individualmente: $d_\infty(x,y) = \max\{|x_1 - y_1|,|x_2 - y_2|\}$, ela também é uma distância válida.

Essa nomenclatura, 1 e ∞, não é ao acaso. Essas distâncias são casos particulares de uma

grande classe de distâncias possíveis em $\mathbb{R}^2$, as distâncias d_p, que são da forma:

$$d_p(x,y) = (|x_1 - y_1|^p + |x_2 - y_2|^p)^{1/p}$$

É fácil ver que o pitágoras é d_2 e que, fazendo $p \to \infty$, teremos que apenas a maior das distâncias sobrevive. No $\mathbb{R}$, todas as distâncias d_p coincidem, mas elas são radicalmente diferentes em dimensões superiores.

E no caso das funções a valores reais? Aqui estamos em uma situação bem mais complicada. Um jeito muito natural de definir distâncias entre duas funções é considerar a área entre elas.

Conseguimos esse valor usando a integral da diferença entre o módulo delas. No caso do meu desenho, eu quero apenas estudar a distância entre funções de $-\pi$ a π para saber qual dentre essas duas, $x - x^3/6$ ou $x - x^3/6 + x^5/120$, se aproxima melhor da função seno (azul entre as duas). Nesse caso, minha distância seria

$$d(f,g) = \int_{-\pi}^{\pi} |f(x) - g(x)|dx$$

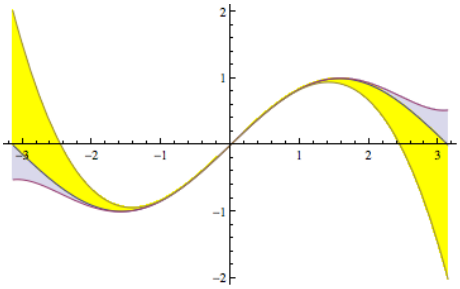


Figura 3: Distância entre funções

Mas você vai reparar um problema sério nessa minha definição. Enquanto eu satisfaço as propriedades 2 e 3 da distância, a reciprocidade e a desigualdade triangular, eu tenho um sério problema na primeira. Se por um lado é verdade que $d(f,f) = 0$, funções iguais possuem integral da diferença nula, não é verdade que a distância é apenas zero para funções completamente iguais. Ora, se eu pego um único ponto de f e mudo, definindo uma f', eu ainda terei que a integral da diferença entre as duas é zero, pois um ponto não interfere na integral, ainda que elas são sejam iguais.

Aqui eu tenho duas opções, mas ambas levam a águas mais profundas a que não me arriscarei nesse texto. A primeira é abandonar a hipótese $d(x,y) = 0 \implies x = y$, o que me daria uma chamada "distância semi-definida" e me forçaria a rever todos os teoremas em que essa hipótese foi usada (unicidade do limite, por exemplo). A segunda seria mudar o espaço de funções. Ao invés de pegar todas as funções, eu posso pegar um novo espaço em que funções que diferem entre si por um número finito de pontos são a mesma. Com esse novo espaço, a integral de uma função só é zero se ela é sempre zero, ou se difere de zero por um conjunto finito de pontos[1] então, em meu espaço, ela é igual a zero. Não me arrisco explicar esse espaço com cuidado (chamado de espaço L^1), mas é legal saber que ele existe e é a opção favorita dos matemáticos

[1]Para ser mais exato, não precisa ser um conjunto finito. Mais para frente você aprende que precisa ser um conjunto *de medida nula*, ou seja, nesse caso, cuja integral é zero. A definição de medida de um conjunto e da exatidão desse conceito é com um curso sério de teoria da medida

frente a esse dilema.

4 Conjuntos abertos e fechados

4.1 Definições

Sofri com o dilema entre apresentar isso antes ou depois de apresentar os conceitos de norma e produto interno, mas como ele não exige a noção de norma eu coloco ele na frente. Vou introduzir aqui conceitos bem abstratos e bem legais, também ferramentas de base para qualquer coisa em física ou matemática, chamados conjuntos abertos. Muitos teoremas são apenas válidos em conjuntos abertos, então é importante saber reconhecer um e conhecer bem suas propriedades para, quando houver um teorema que exija um conjunto aberto, entender a razão dessa exigência.

Antes, vamos ver o conceito de "bola aberta". No resto desse capítulo, X será um conjunto munido de uma distância d, ele é um cara cujos elementos me permitem definir algo como uma distância. A ideia de bola aberta é definir um centro da bola, que chamarei de a, e definir um subconjunto de X composto de todos os caras que distam menos que r de a. r é chamado o raio da bola. O conceito é bem intuitivo, o nome "bola" vem na analogia com o $\mathbb{R}^3$, uma bola aberta em $\mathbb{R}^3$ é de fato uma bola, o conjunto dos pontos que não estão a uma distância maior que r de seu centro a. Uma bola aberta de raio r e centro a é denotada $B(a,r)$, ou $B_r(a)$, vou preferir a segunda notação.

Definição 6 (Bola aberta)**.** *Seja (X,d) um espaço métrico. Seja $a \in X$ e $r \in \mathbb{R}^+$. Definimos uma bola aberta de raio r e centro a, denotada $B_r(a) \subset X$, como:*

$$B_r(a) = \{x \in X \mid d(x,a) < r\}$$

No caso dos reais com a distância do módulo da diferença, uma bola aberta é simplesmente um intervalo aberto (a,b) (ou $]a,b[$ na notação do colegial, que deve ser esquecida). No $\mathbb{R}^2$, a forma de uma bola aberta dependera da distância que você escolher. Vamos ver as bolas abertas das três distâncias apresentadas nos exemplos, d_1, d_2 e d_∞.

Indo um pouco além, é impossível representar uma bola aberta no espaço de matrizes ou no espaço de funções, não se pode desenhar esse tipo de coisa. Ainda, é importante obter essa imagem mental da bola aberta, ela ajuda bastante na compreensão de propriedades e teoremas em espaços e assuntos a que nossa imaginação não chega.

Todo ponto $a \in X$ em que é possível definir uma bola aberta que esteja inteira contida em X é dito um **ponto interior**. De maneira rigorosa, teremos:

Definição 7 (Ponto interior)**.** *Seja (X,d) um espaço métrico. Seja $x \in X$. Dizemos que x é*

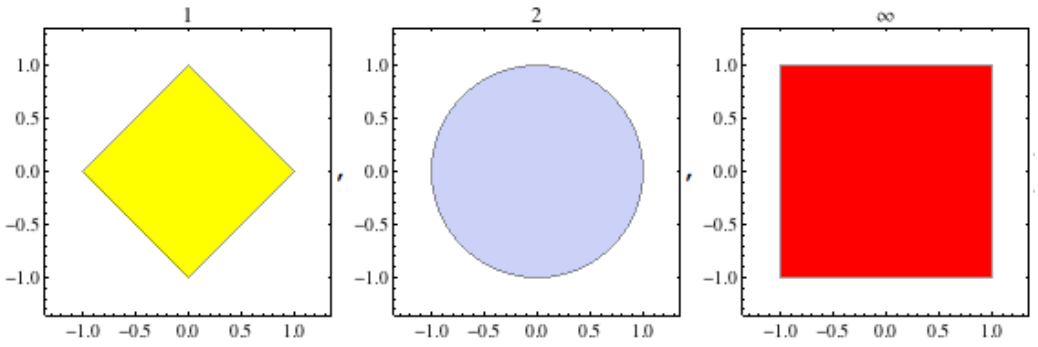


Figura 4: Bolas abertas em diferentes distâncias.

ponto interior de X *se* $\exists \varepsilon > 0 \mid B_\varepsilon(x) \subset X$.

Intuitivamente, é alguém que está "dentro" do conjunto, tão dentro que você consegue definir um conjunto "em volta" dele de caras que estão no conjunto. Se você pegar o $\mathbb{R}$ com a norma da diferença do módulo, vai perceber, por exemplo, que 1 é um ponto interior do conjunto [0,2], enquanto 0 e 2 não são. Note que você consegue definir a bola aberta $B_{1/2}(1)$ em torno de 1 e ela estará inteirinha no [0,2]. No entanto, nenhuma bola aberta centrada em 0 estará inteira contida no [0,2], você sempre terá alguém negativo nessa bola, não há como.

Um ponto interior de um conjunto é algo bem importante e nem sempre é óbvio. Nos reais eles parecem fáceis de detectar, são pontos que não estão na borda do conjunto, literalmente interiores. Nos outros espaços ou outras métricas, nem tanto. A matriz $\left(\begin{smallmatrix} 2 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ é ponto interior do conjunto de matrizes inversíveis 2×2, mas a demonstração disso fica para outro dia. Não consigo demonstrar esse fato sem usar resultados bem mais poderosos, que ainda não citei, é o tipo de caso em que, para provar para uma matriz, é mais difícil que provar para todas.

Citei o exemplo dessa matriz porque o conjunto de matrizes inversíveis 2×2 é chamado $GL_2(\mathbb{R})$, são aquelas que possuem determinante não nulo. Elas são um subconjunto do espaço de matrizes 2×2, chamado $M_2(\mathbb{R})$. É um teorema bem interessante provar que $GL_2(\mathbb{R})$ é um conjunto dito **aberto** em $M_2(\mathbb{R})$, sendo aberto um conjunto em que todos os seus pontos são pontos interiores. Por isso peguei essa matriz, poderia ter pegado qualquer uma com determinante não nulo.

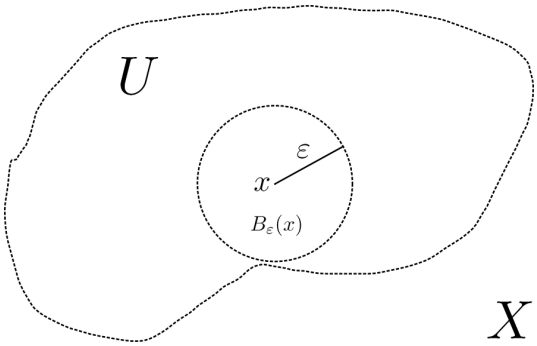


Figura 5: Representação de um conjunto aberto.

Definição 8 (Conjunto aberto)**.** *Seja* (X,d) *um espaço métrico, seja* $A \subset X$*. Dizemos que* A *é um conjunto aberto se todos os seus pontos são pontos interiores.*

É sutil a inclusão dessa ideia de $A \subset X$, um conjunto aberto precisa de um espaço maior

em que ele está inserido para que essa definição tenha alguma utilidade. Nos exemplos que dei, sempre havia um subespaço maior subentendido: $(0,2) \subset \mathbb{R}$ e $GL_2(\mathbb{R}) \subset M_2(\mathbb{R})$. Se você disser "dado X espaço métrico, ele é aberto?", a pergunta em seguida é: "em relação a quem?". Se é em relação a si mesmo ele é sempre aberto, ou seja, se você considerar X seu espaço todo, todos os pontos terão apenas gente de X em volta deles. O espaço todo, então é sempre um aberto.

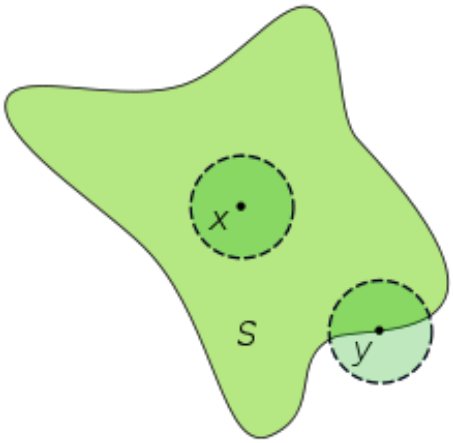


Figura 6: Representação de ponto interior x e ponto de acumulação y.

Intuitivamente, um conjunto é aberto quando, a partir de qualquer um de seus pontos, você pode andar "um pouquinho" em qualquer direção e ainda ficar dentro do conjunto.

O conjunto vazio $\emptyset$ também é um aberto, mas por uma razão meio tensa. Ele satisfaz a definição de aberto: todos os elementos de $\emptyset$ são pontos interiores. Você pode dizer: "mas ele não possui elementos!", e terá razão, mas isso não fere a veracidade da minha afirmação: todos os elementos dele são pontos interiores. É como dizer: "todos os países em que há cavalos voadores são socialistas". Ainda que nenhum país possua cavalos voadores, minha afirmação não pode ser negada, ela é logicamente uma verdade. Assim, o conjunto vazio é um aberto.

Os conjuntos abertos e sua definição possuem uma importante relação com outro tipo de ponto. Falamos sobre o que era um ponto interior, um ponto completamente cercado por elementos do conjunto. Há um outro tipo importante de ponto, que é um ponto "conectado" ao conjunto. Tomemos o conjunto: $(1,2) \cup \{3\}$. Ainda que 1 e 2 não sejam pontos interiores, eles são bem importantes para esse conjunto, e intuitivamente estão mais "conectados" ao conjunto que $\{3\}$. Mas como definir isso de maneira precisa?

Vamos falar que um ponto está "conectado" ao conjunto quando somos capazes de definir uma sequência de caras que tende a ele. O legal é que esse cara não precisa estar no conjunto para ser esse tipo de ponto, dito "ponto de acumulação" do conjunto. Por exemplo, no $(1,2) \cup \{3\}$, 1 e 2 são pontos de acumulação dele, ainda que não estejam nele. São possíveis limites de sequências de $(1,2) \cup \{3\}$.

Definição 9 (Ponto de acumulação)**.** *Seja X espaço métrico e $A \subset X$. Dizemos que $x \in X$ é um ponto de acumulação de A se é possível definir uma sequência inteira diferente de x contida em A cujo limite é x. Em outras palavras, quando:*

$$\exists \{a_n\}_{n \in \mathbb{N}} \subset A \mid \lim a_n = x \in X$$

Em outras palavras, toda bola aberta centrada em x contém ao menos um elemento de A.

Uma outra definição fundamental, cujo poder será relevado mais tarde, é a de conjunto fechado.

Definição 10 (Conjunto fechado)**.** *Seja (X,d) um espaço métrico, seja $F \subset X$. Dizemos que F é um conjunto fechado em X se seu complementar em X, $F^c = X - F = A$, for um conjunto aberto.*

A ideia do conjunto fechado é a de representar um conjunto que contém sua própria "fronteira". Se o complementar dele é um aberto, isso significa que todos os que estão fora dele, no aberto, podem andar um pouquinho e ainda ficar fora dele, ou seja, sempre há uma "separação" entre os que estão fora dele e ele. Assim, ninguém está na "fronteira" dele, pois alguém na fronteira, andando uma distância muito pequena em algum direção, entraria nele.

As aplicações dos conjuntos fechados são bem variadas, vários teoremas funcionam apenas em fechados e eles possuem propriedades interessantíssimas que nos salvam em diversas situações.

4.2 Resultados

O teorema a seguir é tão importante que é uma definição alternativa de conjunto fechado. Ele não somente é importante, mas é um de meus teoremas favoritos, pois ele prova uma relação direta entre dois conceitos que, na definição, não se parecem muito. Pontos de acumulação, na verdade, tem tudo a ver com conjuntos fechados.

Teorema 3. *Um conjunto é fechado se, e somente se, ele contém todos os seus pontos de acumulação.*

Demonstração. Esse "se, e somente se" não está no teorema por acaso. Ele quer dizer que um conjunto fechado contém seus pontos de acumulação, mas também quer dizer que um conjunto só é fechado se contiver todos os seus pontos de acumulação. Ou seja, não somente todos os fechados contém seus pontos, mas todos os que contém seus pontos são fechados. Vamos provar então em duas partes. Primeiro vamos supor que ele é fechado, então provamos que ele contém seus pontos de acumulação (essa demonstração é conhecida como a "ida", denotada $\Rightarrow$). Depois, supomos que ele contém todos os seus pontos de acumulação, então provamos que ele é um fechado (essa demonstração é conhecida como a "volta", denotada $\Leftarrow$).

($\Rightarrow$) Se F é fechado, $F^c = A$ é um aberto. Vamos provar que se x é ponto de acumulação de F, então ele está em F. Ora, basta notar que, se não fosse esse o caso, então teríamos $x \in A$, que é todo mundo que não pertence a F. Mas se $x \in A$ e A é um aberto, é possível achar uma bola centrada em x inteirinha contida em A, com ninguém em F. Ora, isso contraria a

própria definição de ponto de acumulação; então x não seria ponto de acumulação de F. A única possibilidade é, então, que $x \in F$.

($\Leftarrow$) Se F contém todos os seus pontos de acumulação, então não há pontos de acumulação de F em A. Oras, isso significa que nenhum ponto y de A possuirá a propriedade "qualquer bola centrada em y possuirá um elemento de F", pois se não ele seria um ponto de acumulação de F. Sem essa propriedade, isso significa que é possível achar uma bola aberta centrada em y que não contenha ninguém de F, ora, então ela terá apenas gente de A, o que prova que A é um aberto. Disso, deduzimos que seu complementar, F, é fechado. □

Com esse tipo de relação, ou se, e somente se, temos uma equivalência entre essas duas ideias: complementares de abertos e conjuntos que possuem seus pontos de acumulação. Poderíamos até ter usado o fato de conter todos os seus pontos de acumulação como uma definição e fechado e depois provar que seu complementar é sempre um aberto.

Operações com conjuntos são também importantes, e é bom notar que a noção de aberto e fechado é estável por união e intersecção de um número finito de conjuntos.

Teorema 4. *Sejam A_1,$A_2 \subset X$ abertos em X, então $A_1 \cup A_2$ é um aberto em X.*

Demonstração. É mais simples do que parece. Se um $x \in A_1 \cup A_2$, então ele está ou em A_1, ou em A_2, ou nos dois. Em qualquer um desses casos, você pode usar a bola aberta a que tem direito pelo fato de tanto A_1 quanto A_2 serem abertos. Se $x \in A_1$, a bola aberta centrada em x que está inteira contida em A_1 também estará contida completamente em $A_1 \cup A_2$, e isso prova que a união de dois abertos é um aberto. □

Confesso que quando você estiver fazendo suas próprias demonstrações, não escreverá tanto, isso tomaria tempo demais. Precisa apenas deixar suas ideias claras, mas a demonstração deve ser menos contada e mais explícita, essa maneira quase romanceada de demonstrar não é muito comum em livros de matemática. A mesma demonstração acima você encontraria em um livro assim:

Demonstração sem romance. Seja $x \in A_1 \cup A_2$. Suponha, s.p.g., $x \in A_1$. Então $\exists B_\varepsilon(x) \subset A_1 \subset A_1 \cup A_2 \implies A_1 \cup A_2$ é aberto. □

Vale notar que s.p.g. significa "sem perda de generalidade", ou seja, tomei o $x \in A_1$, o que parece específico, porque ele poderia não estar no A_1. O s.p.g. serve para indicar que, tomando $x \in A_2$, o resultado teria sido o mesmo.

Já que vimos o caso da união, passemos à intersecção de abertos.

Teorema 5. *Sejam A_1,$A_2 \subset X$ abertos em X, então $A_1 \cap A_2$ é um aberto em X.*

Demonstração. Sabemos que, dado $x \in A_1 \cap A_2$, temos uma bola aberta centrada em x de raio ε_1 que está inteira contida em A_1 (pois A_1 é um aberto) e também possui uma bola aberta centrada em x de raio ε_2 que está inteira contida em A_2 (pois A_2 também é um aberto). Assim, tomando o menor valor dentre esses dois, $\varepsilon = \min\{\varepsilon_1,\varepsilon_2\}$, uma bola centrada em x com esse raio estará inteira em A_1 e inteira em A_2, então é sempre possível definir uma bola aberta centrada em x que esteja contida em $A_1 \cap A_2$, logo $A_1 \cap A_2$ é aberto. □

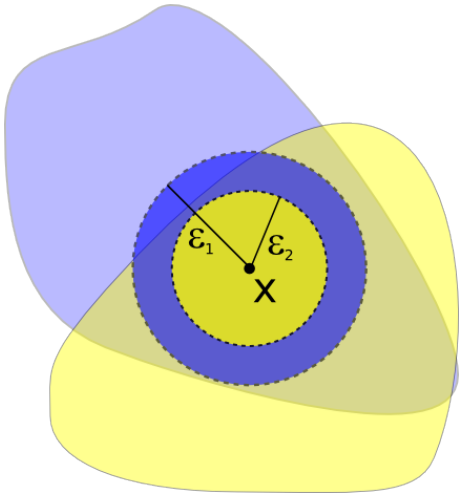


Figura 7: A intersecção de dois abertos é um aberto.

Vale notar uma propriedade importante. Enquanto qualquer união de abertos é um aberto, nem toda intersecção de abertos é aberta. Sim, sei que acabei de provar que dados dois abertos a intersecção dele é aberta, o que é o mesmo que dizer que a interseção de n abertos é aberta, porque posso fazer as intersecções duas a duas e chegar no n. Mas as coisas ficam bem diferentes quando eu mando n para o infinito.

Tomemos o intervalo $(-1/n,1+1/n)$. Conforme o valor de n aumenta, o limite da direita chega mais e mais perto de 1, enquanto o da esquerda vai para zero. Para qualquer valor de n que você escolher, o intervalo $[0,1]$ estará sempre contido nesse intervalo. De fato, se você fizer a intersecção desse conjunto para todos os seus valores de n, terá que o que há em comum entre eles é o intervalo $[0,1]$. De fato, se tentar aumentar um pouco esse conjunto, colocar um $[0,1+\varepsilon)$, por exemplo, encontrará um valor de n, no caso $n > 1/\varepsilon$, tal que $(-1/n,1+1/n)$ não contém o valor $1+\varepsilon$ e, por isso, $1+\varepsilon$ não está na intersecção infinita desses conjuntos. Denotamos:

$$\bigcap_{n\in\mathbb{N}} (-1/n,1+1/n) = [0,1]$$

Assim, a intersecção infinita de abertos não é um aberto. Estaria a demonstração do teorema 5 errada? De forma alguma. Mas note que eu não provei o caso infinito, eu provei só o finito e usei esse fato quando tomei $\varepsilon = \min\{\varepsilon_1,\varepsilon_2\}$. Ora, fosse um número infinito de conjuntos eu teria que pegar o menor valor de ε_n de uma coleção infinita de ε's. ninguém me garante que essa coleção possuirá um menor valor. Afinal, qual o menor valor da coleção $1/n$? Dessa forma, a hipótese da finitude é fundamental para a demonstração da intersecção de abertos.

Essa hipótese não foi feita sobre a união, eu apenas escolhia s.p.g. um dos conjuntos para o x estar e usava a bola do conjunto, cuja existência eu tinha garantida por hipótese. De fato, uma união qualquer, finita ou infinita, de abertos é um aberto.

Existe um conjunto de leis em teoria dos conjuntos chamadas leis de De Morgan, elas

relacionam união e intersecção de conjuntos e seus complementares. A demonstração delas não é difícil, elas são quase um resultado lógico imediato. Se um elemento está no complementar da união entre A e B, então ele não está nem em A, nem em B, então ele estará na intersecção entre os que não estão nem em A, nem em B, ou seja, ele estará na intersecção dos complementares. Escrevendo de forma direta, teremos: $(A \cup B)^c = A^c \cap B^c$. Com isso, podemos escrever:

Teorema 6. *Sejam A_1,$A_2 \subset X$ fechados em X, então $A_1 \cap A_2$ é um fechado em X.*

Demonstração. Se A_1 e A_2 são fechados, seus complementares são abertos. Como $(A^c)^c = A$, podemos usar a lei de De Morgan para escrever $A_1 \cap A_2 = (A_1^c \cup A_2^c)^c$., eu apenas troquei o A pelo A^c na lei. Como A_1^c e A_2^c são abertos, a união deles é aberta e o complementar dela é fechado, então $(A_1^c \cup A_2^c)^c$ é um fechado, bem como $A_1 \cap A_2$. □

Como qualquer união de abertos é aberta, qualquer intersecção de fechados é fechada. A união de fechados, no entanto, possui a mesma limitação da intersecção de abertos: ela vale apenas para um número finito de operações. A união de todos os $[1/n, 1 - 1/n]$ dá o aberto (0,1). Uma demonstração de que união finita de fechados é fechada segue a mesma linha das anteriores. Deixo como exercício, porque não há um só texto de análise que não te deixe algo de presente como exercício.

É importante notar que fechado não é o contrário de aberto. O vazio, por exemplo, é aberto, pois todo ponto dele contém uma bola com apenas elementos dele. Mas ele também é fechado, pois contém todos os seus pontos de acumulação. Você pode dizer: "mas ele não tem nenhum ponto de acumulação!", e terá razão, mas também há de concordar que todos os que ele tem estão dentro dele. Assim, o vazio é fechado e aberto. O espaço todo também é considerado fechado e aberto em si mesmo, pois você jamais consegue sair dele com uma bola aberta em qualquer ponto (ele é tudo o que existe) e o complementar dele é o vazio, que é aberto.

5 Funções contínuas

5.1 Definições

A noção de função contínua é fundamental em qualquer área da matemática ou da física, é difícil exagerar na importância desse conceito. Uma função contínua pode ser informalmente definida como uma função que leva pontos próximos em pontos próximos, sem saltos, sem surpresas.

Não é estranho, portanto, definir uma função contínua da seguinte forma:

Definição 11 (Função contínua)**.** *Diz-se de $f : X \to Y$ uma função contínua se ela coincide com o valor de seu limite para todos os pontos de seu domínio.*

Essa definição é o mesmo que reescrever a definição de limite trocando o L pelo valor da função de a:

$$\forall \varepsilon > 0, \forall a \in X, \exists \delta > 0 \mid d_X(x,a) < \delta \implies d_Y(f(x),f(a)) < \varepsilon$$

5.2 Resultados

Teorema 7. *Sejam X e Y espaços métricos e $f : X \to Y$ uma função contínua. Sejam x_n e y_n sequências em X e Y respectivamente tais que $y_n = f(x_n)$. Se $x_n \to x$ e $f(x) = y$, então $y_n \to y$. Ou seja, funções contínuas levam sequências convergentes em sequências convergentes, respeitando a atribuição do limite.*

O grande poder desse teorema é a manipulação do limite para dentro e fora da função. Em geral, $\lim f(a_n) \neq f(\lim a_n)$, esse teorema me garante que, no caso das funções contínuas, nós teremos a igualdade. Esse tipo de situação aparece bastante, é talvez a propriedade mais útil das funções contínuas.

Demonstração. Se $x_n \to x$, podemos escrever a definição da convergência dessa sequência: para qualquer ε_x existe um N tal que, a partir dele, $d_X(x_n,x) < \varepsilon_x$. Como $f(x_n) = y_n$ e f é uma função contínua, eu escrevo a definição de função contínua: dado um $\varepsilon > 0$, existe um δ tal que $d_X(x_n,x) < \delta \implies d_Y(f(x_n),f(x)) < \varepsilon$.

A dica é colocar essas duas definições juntas. Na definição da função contínua, você precisa de um δ para cada $\varepsilon > 0$, ou seja, precisa colocar o x_n e o x próximos de uma distância δ. Ora, essa sequência converge, então você consegue achar um N tal que, a partir dele, x_n e x estejam próximos de uma distância δ. Assim:

$$\forall \varepsilon > 0, \exists N \mid n > N \overset{1}{\implies} d_X(x_n,x) < \delta \overset{2}{\implies} d_Y(f(x_n),f(x)) < \varepsilon \implies d_Y(y_n,y) < \varepsilon$$

Note que a implicação 1 é a definição da convergência de x_n, a implicação 2 é a definição da continuidade da função f e o resultado final, se você esquece o meio, é a convergência da sequência $y_n \to y = f(x)$. □

Como um resultado legal e interessante, vou misturar os conceitos vistos até agora para provar uma propriedade interessante da análise do espaço de matrizes. Já falamos sobre o espaço $GL_n(\mathbb{R})$, o espaço das matrizes reais $n \times n$ que são inversíveis, possuem o determinante não nulo. A propriedade é a seguinte:

Teorema 8. *$GL_n(\mathbb{R}) \subset M_n(\mathbb{R})$ é um conjunto aberto.*

Demonstração. Não vou provar que $GL_n(\mathbb{R})$ é aberto, confesso que não sei fazer isso diretamente, vou provar que o complementar dele é fechado. Se $GL_n(\mathbb{R})$ são as matrizes inversíveis,

então o complementar delas são as não inversíveis, também chamadas de singulares, as matrizes $n \times n$ que possuem determinante zero. Para provar que esse conjunto é fechado, vou provar que ele contém todos os seus pontos de acumulação.

Seja A_n uma sequência de matrizes singulares que converge, vamos provar que o limite delas também é singular, isso provaria que o conjunto das matrizes singulares é fechado (pois contém qualquer limite de suas sequências, seus pontos de acumulação). Aplicando o determinante nas matrizes, teremos uma sequência de determinantes: $\det A_1, \det A_2 \ldots$, e sabemos que essa sequencia é constante igual a zero, então seu limite é zero: $\lim(\det A_n)) = 0$. Como o determinante é uma função contínua (não vou provar isso, está implícito na definição de determinante), podemos passar o limite para "dentro" dele, pelo teorema que provamos acima, teremos: $\det(\lim A_n) = 0$, então a matriz limite também é singular, o conjunto é fechado e seu complementar, $GL_n(\mathbb{R})$, é aberto. □

6 Espaços completos

6.1 Definições

A noção de espaço completo é igualmente importante na análise, muitos teoremas são apenas válidos quando temos a garantia da completude, que nem sempre é assegurada. Essa propriedade, aliás, é a grande diferença entre os reais e os racionais, como nós vamos ver. Para explicar melhor isso, vamos introduzir uma ideia diferente de convergência de sequências.

Ao invés de falar que uma sequência converge se ela vai ficando cada vez mais próxima de um valor, vamos pensar na convergência de sequências como os elementos delas ficam cada vez mais próximos uns dos outros. Pode parecer a mesma coisa, mas se deve tomar cuidado, ninguém garante que os elementos ficarem mais e mais próximos levará a sequência para algum lugar. Na verdade, não é bom você dizer que uma sequência com termos ficando mais e mais próximos converge, por isso, damos outro nome para esse tipo de sequência:

Definição 12 (Sequência de Cauchy)**.** *Seja X um espaço métrico e $\{a_n\} \subset X$ uma sequência. Dizemos que $\{a_n\}$ é uma sequência de Cauchy se:*

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \mid n,m > N \implies d(a_n,a_m) < \varepsilon$$

Ou seja, sempre conseguimos encontrar um N grande o suficiente para que, a partir dele, todos os elementos estejam bem próximos entre si, essa é a ideia da sequência de Cauchy.

Eis a pergunta: toda sequência de Cauchy converge? A resposta não é fácil, e é a chave da definição de conjunto completo. Digo, desde já, que não, a convergência das sequências de Cauchy não é garantida em geral. Nos espaços métricos em que todas as sequências de Cauchy convergem, dizemos que todas essas sequências "tem para onde ir" quando seus elementos ficam

indefinidamente próximos. Esses conjuntos possuem todos os limites de suas sequências de Cauchy, nada mais justos que os chamar de conjuntos completos:

Definição 13 (Espaço completo)**.** *Seja X um espaço métrico. Dizemos que X é um espaço completo se toda sequência de Cauchy em X possuir limite em X.*

6.2 Resultados

Como os personagens desse capítulo já foram apresentados, vamos entender como eles se relacionam entre si e com o resto dos que já conhecemos.

Em primeiro lugar, vamos dar exemplos de sequências de Cauchy que não convergem. Para isso, precisamos encontrar um espaço não-completo, mas não é preciso ir muito longe, os racionais já não satisfazem esse critério. Por quê? Imagine a sequência dos elementos cujo quadrado fica mais e mais próximo de 2, e a diferença entre os elementos da sequência são as casas decimais. Os primeiros elementos da sequência são: $1, 1{,}4, 1{,}41, 1{,}414, 1{,}4142, 1{,}41421, \ldots$. Eles são uma sequência crescente tal que $a_n^2 < 2$. Como ele sempre ganha uma casa decimal, a diferença entre o a_n e o a_{n+1} deve sempre ser menor que 10^{-n}, que é a precisão decimal que ele ganha.

É claro que essa sequência, se estivesse nos reais, convergiria para $\sqrt{2}$. Mas ela não está nos reais. No entanto, a diferença entre os elementos diminui rapidamente com o passar dos elementos. Se eu escolher um $\varepsilon > 0$, sempre posso achar um N tal que $10^{-N} < \varepsilon$ e todos depois dele estarão distantes entre si de uma distância menor que ε, oras, então essa sequência é uma sequência de Cauchy. No entanto, não existe nenhum $L \in \mathbb{Q}$ que satisfaça minha condição de convergência. Essa sequência de Cauchy não converge em $\mathbb{Q}$, ainda que todos os seus elementos sejam racionais. Assim, os racionais não são um espaço completo.

Por outro lado, toda sequência que converge é de Cauchy. Isso é bem intuitivo, se os elementos de uma sequência vão ficando infinitamente próximos de um elemento L, então eles vão ficando cada vez mais próximos entre si. Naturalmente, provamos o resultado:

Teorema 9. *Toda sequência convergente é de Cauchy.*

Demonstração. Sejam X espaço métrico e $\{a_n\} \subset X$ uma sequência convergente. Vamos provar que $\{a_n\}$ é de Cauchy. Ora, se $\{a_n\}$ converge para $L \in X$, então, para um $\varepsilon > 0$, sempre podemos encontrar um N, por maior que ele precise ser, tal que $n > N \implies d(a_n,L) < \varepsilon$. Ora, se tomamos quaisquer $n,m > N$, teremos que:

$$d(a_n, a_m) \leq d(a_n,L) + d(a_m,L) \leq \varepsilon + \varepsilon = 2\varepsilon$$

Usei apenas a desigualdade triangular e a simetria da função distância ($d(x,y) = d(y,x)$). Isso prova que os elementos se aproximam indefinidamente, que a sequência é de Cauchy. □

É importante comentar, os reais são um conjunto completo, mas uma demonstração disso exigiria o uso de propriedades dos reais e, para isso, eu precisaria te contar como os reais são definidos. Na verdade, alguns autores preferem colocar a propriedade da completude como um axioma dos reais, então eles são completos por definição.

Teorema 10. *Sejam X espaço métrico e $C \subset X$ um conjunto completo. Então C é fechado em X.*

Demonstração. Esse teorema é quase reescrever a definição de conjunto completo. Se toda sequência de Cauchy converge em C, então toda sequência convergente (que também é de Cauchy, pelo teorema acima), converge em C, então C contém todos os seus pontos de acumulação e, pelo teorema 3, é fechado. □

Há outro teorema importante envolvendo completos, ele é uma poderosa ferramenta de demonstração de outros resultados. Se aplicado aos números reais, ele é conhecido como o teorema dos intervalos encaixantes, o nome lhe é próprio. Mas ele não precisa ser restrito aos reais não, serve em qualquer espaço métrico, como vamos ver:

Teorema 11 (da intersecção de Cantor)**.** *Seja $\{F_n\}$ uma sequência de conjuntos fechados, completos, não-vazios e encaixados, ou seja, $F_{n+1} \subset F_n$. Suponha que o diâmetro de F_n tenda a zero. Então a intersecção de todos os F_n não é vazia e é composta de apenas um ponto.*

Demonstração. Eis o teorema: a família F_n é como uma dessas bonecas russas matryoshka, um conjunto dentro do outro dentro do outro com o tamanho diminuindo a cada passagem, tendendo a zero no infinito. Vamos tomar a sequência x_n sendo x_n alguém pertencente a F_n (o que posso fazer, porque os conjuntos não são vazios). Como $F_{n+1} \subset F_n$, então toda a sequência a partir de x_n está em F_n. Mas note que $d(x_n,x_m) < \text{diam}(F_n)$, porque o diâmetro é a maior distância autorizada entre dois elementos do conjunto. Mas conforme vamos avançando na sequência, o diâmetro do conjunto em que os pontos estão vai diminuindo e isso vai empurrando $d(x_n,x_m)$ para zero (teorema do confronto).

Ora, se isso acontece, então x_n é uma sequência de Cauchy. Como F_n é completo, existe um x que é o limite de x_n. Como F_n é fechado e x é um ponto de acumulação seu (teorema 3), teremos que $x \in F_n$. Como eu poderia ter escolhido o n que eu bem entendesse, esse x estará em todos os conjuntos da família, então, tomando a intersecção de todos da família, teremos o x na intersecção.

Seria ele o único? Se não fosse, se existisse um $y \neq x$ também na interseção de todos, eu teria $d(x,y) = \varepsilon > 0$. Mas eu sempre posso achar um n grande o suficiente para que $\text{diam}(F_n) < \varepsilon$. Ora, com isso eu estaria dizendo que dois caras que estão em todos os F_n não podem estar em um deles, pois a distância entre eles seria maior que o diâmetro do conjunto, que é a maior distância possível entre dois elementos. Como isso é absurdo, o x é o único elemento da intersecção. □

Um teorema, um pouco mais avançado, convenceu-me a ser colocado nesse texto. Ele tem uma demonstração meio espinhosa, mas é um resultado lindo e tem aplicações diretas e práticas em qualquer área das exatas. Antes de enunciá-lo, preciso te contar o que é uma função contractante. Intuitivamente, é uma função que leva dois pontos a outros dois pontos mais próximos, é como se sua imagem estivesse agora mais "densa" que seu domínio.

Definição 14 (Função contractante)**.** *Sejam X espaço métrico e $f : X \to X$. Dizemos que f é contractante se, para todo x,$y \in X$:*

$$d(f(x),f(y)) \leq q.d(x,y),$$

com $q < 1$. Nesse caso, dizemos que a função é q-contractante.

Corolário 2. *Toda aplicação contractante é contínua.*

Coloquei esse cara como corolário pois nem vale uma demonstração. Se você quer deixar $d(f(x),f(y)) < \varepsilon$, basta tomar $d(x,y) < \varepsilon$ que isso já garante que $d(f(x),f(y)) < q.d(x,y) < q.\varepsilon < \varepsilon$.

Essa definição serve para o grande teorema que segue. Não somente ele é poderoso e importante, mas a demonstração possui beleza particular:

Teorema 12 (do ponto fixo de Banach)**.** *Seja X um espaço métrico completo e $f : X \to X$ uma função contractante. Então a função f possui um único ponto fixo, ou seja, existe um único $x \in X$ tal que $f(x) = x$.*

Demonstração. Meu objetivo é encontrar um x^* tal que $f(x^*) = x^*$. Vamos começar escolhendo um $x_0 \in X$ e definindo uma sequência: $x_n = f(x_{n-1})$, com o x_0 de elemento inicial. Estudemos como os elementos dessa sequência se afastam uns dos outros, vamos pegar um N bem grande e tomemos dois n,$m > N$. Sem perda de generalidade, vamos supor $n > m$, e vejamos como se comporta $d(x_n,x_m)$. Aplicando várias desigualdades triangulares, eu posso escrever:

$$d(x_n,x_m) \leq d(x_n,x_{n-1}) + d(x_{n-1},x_{n-2}) + \ldots + d(x_{m+1},x_m).$$

Mas é fácil ver que todo $d(x_n,x_{n-1})$ é menor que $d(x_1,x_0)$. Isso porque $d(x_2,x_1) = d(f(x_1),f(x_0)) \leq q.d(x_1,x_0)$, pela própria definição de função contractante. Por indução, dá pra perceber que $d(x_n,x_{n-1}) \leq q^{n-1}d(x_1,x_0)$. Escrevemos isso naquela desigualdade acima:

$$\begin{aligned} d(x_n,x_m) &\leq d(x_n,x_{n-1}) + d(x_{n-1},x_{n-2}) + \ldots + d(x_{m+1},x_m) \\ &\leq q^{n-1}d(x_1,x_0) + q^{n-2}d(x_1,x_0) + \ldots + q^m d(x_1,x_0) \\ &= \left(q^{n-1} + q^{n-2} + \ldots + q^{m-1}\right) d(x_1,x_0) = \sum_{k=m}^{n-1} q^k d(x_1,x_0). \end{aligned}$$

Então eu sei que a distância entre dois caras da sequência cujos índices estão depois de N possuem entre si uma distância menor que aquele termo com o somatório. Nesse termo, vou deixar um q^m em evidência e reescrever os limites: $\sum_{k=m}^{n} q^k d(x_1,x_0) = q^m \left(\sum_{k=0}^{n-m-1} q^k\right) d(x_1,x_0)$. Por enquanto isso é só uma manipulação para chegar onde quero. Aquele termo da somatória pode ter o valor que quiser, mas certamente será menor que se eu somar, ao invés de 0 a $n-m-1$, de 0 a ∞. Por isso, escrevo:

$$\begin{aligned} d(x_n,x_m) &\leq \sum_{k=m}^{n-1} q^k d(x_1,x_0) \leq q^m \left(\sum_{k=0}^{n-m-1} q^k\right) d(x_1,x_0) \\ &\leq q^m \left(\sum_{k=0}^{\infty} q^k\right) d(x_1,x_0) = q^m \frac{1}{1-q} d(x_1,x_0). \end{aligned}$$

Eis a magia. Se eu pegar um $\varepsilon > 0$, por menor que seja, eu sempre consigo encontrar um N tal que $q^N < \varepsilon \frac{1-q}{d(x_1,x_0)}$. Isso vai forçar aquele termo à direita a ficar menor que ε, para todo $n,m > N$. Ora, eu provei que, dado $\varepsilon > 0$, existe N tal que $n,m > N \implies d(x_n,x_m) < \varepsilon$, então a sequência que eu defini é de Cauchy.

Como o espaço métrico X é completo (por hipótese), essa sequência converge para um limite x^*.

Se $f(x_{n-1}) = x_n$, vamos tomar o limite dos dois lados dessa expressão. Teremos $\lim f(x_{n-1}) = \lim x_n = x*$, que é o limite de x_n. Mas toda contractante é contínua (corolário 2), e como você pode passar o limite para "dentro" da função contínua (teorema 7), escrevemos $\lim f(x_{n-1}) = f(\lim x_{n-1}) = f(x^*) = x^*$. Então provamos que f possui um ponto fixo.

Provemos agora que ele é único. Como na nossa prova de unicidade do limite, vamos supor que existe um outro ponto depois, depois provaremos que ele é igual a nosso x^*. Suponha que existe um y tal que $f(y) = y$. Então:

$$0 \leq d(x^*,y) = d(f(x^*),f(y)) \leq q.d(x^*,y).$$

Ora, conhece algum número menor ou igual a si mesmo vezes um número menor que um? Só o zero pode satisfazer isso, então $d(x^*,y) = 0$. Pela primeira propriedade das distâncias, isso implica que $x^* = y$ e o ponto fixo é único. □

Eu confesso que essa não foi uma demonstração fácil. Eu mesmo não seria capaz de a fazer sozinho, sem ajuda, e não lembrava dela antes de checar em um livro os elementos dela. Ainda, as técnicas usadas nela não são estranhas, qualquer função contractante pede uma p.g. infinita, porque a razão menor que um chama a convergência da série.

Eu sei que não é o lugar dele aqui, mas o teorema do ponto fixo de Banach chama um outro teorema muito legal e poderoso, ainda que um pouco avançado. Eu não resisto, ainda que tenha me segurado bastante, a enunciar e provar um dos teoremas mais importantes das

equações diferenciais. Se você quiser ignorar todo o resto dessa seção de resultados, não o culpo, porque esse teorema não tem muito a ver com todo o resto da apresentação, ele só é lindo demais para ser omitido. Antes, preciso te contar o que é uma função lipschitziana.

Definição 15 (Função lipschitziana)**.** *Sejam X e Y espaços métricos. Dizemos que a função $f : X \to Y$ é lipschitziana se, para todo par x,$y \in X$, existe uma constante $M \in \mathbb{R}$ tal que:*

$$d_Y(f(x),f(y)) \leq M.d_X(x,y).$$

Nesse caso, dizemos que a função é M-lipschitziana, onde M é sua constante de Lipschitz.

Corolário 3. *Toda função lipschitziana é contínua.*

Intuitivamente, essa função possui um crescimento controlado, não explode, e qualquer distância na imagem pode ser comparada à distância dos equivalentes no domínio por uma mesma constante. Não é uma propriedade fraca, a função x^2, por exemplo, não é lipschitziana. Nesse caso, eu nunca consigo um único M, por maior que ele seja, que seja capaz de tornar, para todo x e y, verdadeira a desigualdade $|x^2 - y^2| \leq M.|x - y|$.

Se eu conseguir essa propriedade mas se para cada ponto $x \in X$ eu precise de uma constante M diferente, dizemos que a função é **localmente lipschitizana**.

E vamos usar como usar esse conceito para o próximo teorema, que vou passar em uma ordem diferente. Primeiro vou provar, depois enunciar, para aumentar o suspense.

Pegue a seguinte função:

$$G(y) = y_0 + \int_{x_0}^{x_1} f(x, y(x))dx$$

Com $y(x_0) = y_0$. E eu exijo duas propriedades de f:

1. f é contínua em x.

2. f é M-lipschitziana em y.

A propriedade 2 significa que, para todo par de funções y_2 e y_1, $|f(x, y_2(x)) - f(x, y_1(x))| \leq M.d(y_2,y_1)$, sendo d a distância que você preferir definir no espaço de funções. Como y, f e G são funções reais, eu me permito usar a distância do módulo da diferença para elas. Vamos

estudar como se comporta $|G(y_2) - G(y_1)|$:

$$\begin{aligned}
|G(y_2) - G(y_1)| &= \left| \int_{x_0}^{x_1} f(x, y_2(x))dx - \int_{x_0}^{x_1} f(x, y_1(x))dx \right| \\
&= \left| \int_{x_0}^{x_1} f(x, y_2(x))dx - f(x, y_1(x))dx \right| \\
&\overset{1}{\leq} \int_{x_0}^{x_1} |f(x, y_2(x)) - f(x, y_1(x))|dx \\
&\overset{2}{\leq} \int_{x_0}^{x_1} M_1.d(y_2,y_1)dx \\
&\overset{3}{=} M.d(y_2,y_1)|x_1 - x_0|
\end{aligned}$$

Novamente numerei as passagens para explicar com calma. Na 1, eu uso a desigualdade $\left|\int_a^b f(x)dx| \leq \int_a^b |f(x)\right| dx$, que parece bem evidente, pois somar vários caras e depois tirar o módulo dará certamente um resultado menor ou igual a somar todos já positivos (isso te lembra a desigualdade triangular? Pois deveria!). Na 2, eu uso o fato de f ser lipschitziana. na 3, eu percebo que a integral não depende mais de x e realizo a integral, que dá apenas o integrando vezes o intervalo.

Agora vem a manobra que muda o jogo: é verdade que o x_0 está fixo, mas quem escolhe o x_1 sou eu. Assim, posso pegar um x_1 tal que o intervalo $|x_1 - x_0|$ fique tão pequeno a ponto de $M|x_1 - x_0|$ ser menor que um! E para que isso vai servir? Eu escrevo $M|x_1 - x_0| = q < 1$ e terei que, para todo y_2, y_1, teremos: $|G(y_2) - G(y_1)| \leq q.d(y_2.y_1)$, então G é contractante! Assim, se o espaço de funções é um espaço completo com a distância que você definiu, *pelo teorema do ponto fixo de Banach* (teorema 12), existe um intervalo $[x_0,x_1]$ tal que, nesse intervalo, existe um único ponto fixo da função G, ou seja, um único $y*$ tal que:

$$G(y^*) = y^*(x) = y_0 + \int_{x_0}^{x} f(x', y^*(x'))dx'$$

E o que eu ganho com isso? Para entender a dimensão desse resultado, você precia derivar os dois lados dessa equação para chegar a:

$$\frac{dy(x)}{dx} = f(x, y(x))$$

E, finalmente, poder dizer:

Teorema 13 (de Cauchy-Lipschitz)**.** *Seja $f = f(x, y(x))$ uma função contínua em x e lipschitziana em y. Seja $y(x_0) = y_0$. Então existe um intervalo $[x_0, x_1]$ tal que a equação diferencial $\frac{dy(x)}{dx} = f(x, y(x))$ possui solução única.*

Usando alguns gracejos na demonstração, podemos relaxar a hipótese de lipschitziana e

provar que basta ser localmente lipschitziana em y para que a equação diferencial possua solução única em um intervalo.

Um exemplo interessante da aplicação das hipóteses desse teorema é um tirado da física. Se queremos estudar a equação diferencial associada ao escoamento de um líquido em um funil, queremos estudar a variação da altura y com o tempo t, chegaremos a uma equação diferencial da forma $y' = \alpha\sqrt{y}$. Essa função é localmente lipscitziana em todo lugar, exceto na origem (a derivada dela explode na origem). Assim, você terá solução única se você der como condição inicial a função $y(t_0) = y_0$ em qualquer lugar, exceto se der $y(t_0) = 0$. Nesse ponto, o teorema não se aplica.

Intuitivamente, isso é razoável. Se eu te contar que, nesse momento, o funil está vazio (altura nula), você nada pode dizer sobre o passado dele. Ele pode ter esvaziado nesse instante, há dois minutos, ele pode nunca ter estado cheio. Se eu disser que a altura da água são 3cm, você saberá exatamente onde ela estava há alguns segundos e onde estará em poucos instantes, mas a condição inicial da altura nula, ainda que permita uma solução, não garante a unicidade, vários "históricos" da água no funil são possíveis; enquanto em qualquer outra condição inicial apenas um passado e um futuro são aceitos para a água descendo o funil.

7 Análise Real

7.1 Definições

Vamos dedicar algumas linhas a especificar resultados do que provamos para os reais. Pode parecer uma especificação desnecessária, mas não é, já que os reais são um dos mais úteis dos espaços. Os resultados apresentados nessa seção valem apenas para os reais.

E quais as propriedades interessantes do reais? Eles são um corpo[2] completo, ordenado (existe uma noção de "maior que" e "menor que", diferentemente dos complexos, que não possuem uma relação de ordem evidente) e arquimediano (dados um x e y positivos tais que $x > y$, existe um n natural tal que $ny > x$).

Essas duas noções, completo e arquimediano, decorrem de uma um pouco mais forte, chamada de "propriedade do supremo". Definimos em todo conjunto que possui a noção de ordem as noções de supremo e de ínfimo desse conjunto.

Definição 16 (Majorante)**.** *Seja A um conjunto com uma relação de ordem. Dizemos que a é um majorante de A se $x \leq a, \forall x \in A$.*

Note que a não precisa estar no conjunto. Como exemplo, temos que 2 é um majorante do conjunto [0,1]. O supremo não é somente isso. Ele é um majorante, mas um majorante especial.

[2]Definimos um pouco mais para frente, na seção 9.1, definição 21, por enquanto tenha em mente que um corpo é um conjunto com duas operações, multiplicação e soma, que funcionam muito bem.

Definição 17 (Supremo)**.** *Seja A um conjunto com uma relação de ordem. Dizemos que s é o supremo deste conjunto se ele é o menor dos majorantes. Ou seja, se $x \leq s, \forall x \in A$ e se y é um majorante de A, então $s \leq y$.*

As noções de minorante e ínfimo são análogas, mas trocando o sinal das desigualdades. Um minorante é um elemento menor que todos os do conjunto e o ínfimo é o maior dos minorantes. E a existência de um supremo não é garantida. Na verdade, os reais são importante por serem o único corpo ordenado com a propriedade do supremo, ou seja, um conjunto em que o supremo de todo subconjunto limitado superiormente está garantido.

Propriedade do supremo: *Todo subconjunto limitado superiormente de $\mathbb{R}$ possui supremo.*

Isso não é um teorema, é uma forma de definir o conjunto dos reais. Na verdade, é possível provar, com isso que os reais são um corpo completo e arquimediano. Evidentemente, todas essas considerações se aplicam também aos ínfimos, tomando-se conjuntos limitados inferiormente.

7.2 Resultados

Existem alguns resultados fundamentais na análise real, evocarei alguns aqui. Infelizmente deixarei de lados dois gigantes dessa matéria, o teorema do valor médio e o teorema fundamental do cálculo, porque não quero entrar nas águas perigosas da derivação, deixo isso para outro dia. O primeiro resultado é o teorema da convergência de sequências monótonas dominadas. Ele é necessário para a demonstração de um outro resultado importante, o teorema de Bolzano-Weierstrass, mas é também um uso interessante da relação de ordem para forçar a convergência de sequências.

Teorema 14. *Uma sequência real crescente possui limite se, e somente se, é limitada.*

Demonstração. ($\Rightarrow$) Vamos supor que a sequência seja limitada. Provemos que ela possui limite. Se a_n é limitada, então o conjunto $\{a_n\}$ é limitado e está contido em $\mathbb{R}$. Pela propriedade do supremo, existe um c supremo desse conjunto. Provemos que c é o limite da sequência.

Seja $\varepsilon > 0$. É fácil ver que sempre é possível encontrar um N tal que $a_N > c - \varepsilon$, do contrário $c - \varepsilon$ seria um majorante e, sendo menor que c, seria o supremo, o que contradiz a definição de c. Assim, $\forall \varepsilon > 0$, existe um N tal que, a partir dele, $|c - a_n| < \varepsilon$, o que prova que c é o limite da função.

($\Leftarrow$) Se $a_n \to c$, então $c \geq a_n, \forall n$ pois, se houvesse um N tal que $a_N > c$, então todos os elementos da sequência a partir de N seriam maiores que c e estarão sempre distantes de c a uma distância de pelo menos $|c - a_N|$. Logo, para $\varepsilon = |c - a_N|$, a convergência não seria nunca possível. Assim, a_n é limitada e c é um majorante seu.

□

Como prometido, segue o teorema de Bolzano-Weierstrass. Ele possui uma generalização possível, o que quase me fez postergar sua apresentação e mostrar apenas o caso geral, para os reais você apenas aplicaria o teorema para o caso real. Mas ele é tão famoso que merece enunciado e demonstração próprios. A ideia dele é muito simples: você sabe que algumas sequências em $\mathbb{R}$, ainda que limitadas, não convergem. Um exemplo é a $a_n = (-1)^n$. Ela não é maior que 1 nem menor que -1, mas não vai a lugar nenhum no infinito. No entanto, Bolzano-Weierstrass vai te garantir que ela possui uma subsequência convergente, você sempre consegue pescar elementos em uma dada ordem tal que, definindo com eles uma nova sequência, ela possua um limite. Nesse exemplo, os elementos de n par convergem para 1 e os ímpares, para -1.

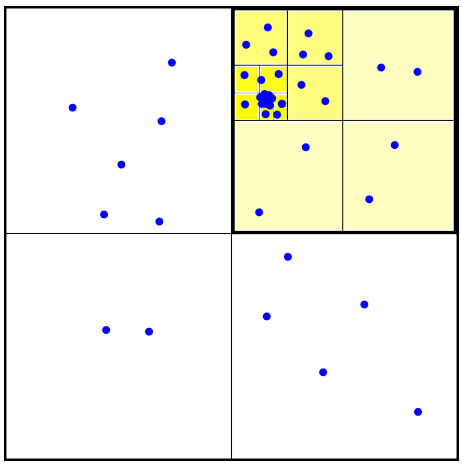


Figura 8: Generalização de Bolzano-Weierstrass para $\mathbb{R}^2$.

Teorema 15 (de Bolzano-Weierstrass)**.** *Toda sequência real limitada possui subsequência convergente.*

Demonstração. Para demonstrar esse teorema, vamos usar uma técnica divertida, que vamos repetir ainda nesse texto para Heine-Borel. Se uma sequência $\{x_n\}$ é limitada em $\mathbb{R}$, podemos colocá-la toda dentro de um intervalo $[a_1,b_1]$. Em seguida, partimos o intervalo em dois, usando o ponto médio para a imagem ficar mais clara: $[a_1,b_1] = [a_1,M_1] \cup [M_1,b_1]$. Não sei muito sobre $\{x_n\}$, mas sei que ela é infinita e, em pelo menos um desses intervalos, há infinitos elementos dela. Vamos pegar um elemento x_{k1} pertencente ao intervalo que contém infinitos.

Em seguida, partimos o intervalo que possui infinitos em mais dois pedaços. Certamente um deles conterá infinitos elementos. De novo, pegamos um x_{k2} da sequência que pertence ao intervalo que possui infinitos elementos.

Já deu pra perceber que vamos fazer isso várias vezes. Na m-ésima vez, teremos um intervalo do tipo $I_m = [a_m,b_m]$ com infinitos elementos. Seu diâmetro será $2^{-m}.(b_1 - a_1)$ e, dentro desse intervalo, eu vou pegar um x_{km}, que existe por construção. Ora, todo intervalo I_m é fechado, limitado, completo e forma uma família encaixante, $I_{m+1} \subset I_m$. Pelo teorema da intersecção de Cantor (teorema 11), teremos que existe um único x que pertence a todos os I_m's. E como cada I_m contém ao menos um elemento da sequência inicial (na verdade contém sempre infinitos, é a definição do I_m), tendo os I_m's diâmetro que tende a zero, eu sempre posso, para todo $\varepsilon > 0$, encontrar um I_N com diâmetro menor que ε, ou seja, $2^{-N}.(b_1 - a_1) < \varepsilon$ (propriedade arquimediana dos reais). E todos os seguintes elementos da sequência que eu defini pegando um cara de cada I_m está a uma distância menor que ε de x. Logo, essa subsequência x_{kn} que eu defini converge e seu limite é x. □

É, eu acabei desistindo de usar o teorema da convergência de sequências monótonas no Bolzano-Weierstrass, preferi usar a intersecção de Cantor, que é bem mais forte e deu conta no recado. Esse resultado pode ser facilmente generalizado para $\mathbb{R}^n$. Coloquei uma figura para dar uma ideia de como seria essa demonstração no $\mathbb{R}^2$.

Nos reais, você dificilmente vai ouvir falar do teorema da intersecção de Cantor, aqui ele ganha o nome de "teorema dos intervalos encaixantes", mas é somente um caso particular daquele primeiro teorema, bem mais poderoso.

Em seguida, chamo talvez o resultado mais importante dessa seção. Ele é importante na medida em que não possui generalização trivial, pois ele usa a noção de ordem diretamente, ele fala de "todos os valores entre $f(a)$ e $f(b)$". Nos complexos, não é possível falar em valores entre $2+i$ e $4-3i$.

Teorema 16 (do valor intermediário)**.** *Seja $f : [a,b] \to \mathbb{R}$ uma função contínua. Então f assume todos os valores entre $f(a)$ e $f(b)$.*

Demonstração. Sem perda de generalidade, vamos supor que $f(a) < f(b)$. Em outras palavras, eu digo que, se tomarmos um u tal que $f(a) \leq u \leq f(b)$, então existe um $c \in [a,b]$ tal que $f(c) = u$.

Tomemos o conjunto S dos elementos $x \in [a,b]$ tais que $f(x) \leq u$. Ele é um subconjunto de $[a,b]$ e não é vazio, já que pelo menos a está nele. Pela propriedade do supremo, existe um $\sup S$ que é o menor dos majorantes de S. Seja $c = \sup S$, vamos provar que $f(c) = u$.

Imagine que não seja esse o caso. Então ou $f(c) > u$, ou $f(c) < u$. No primeiro caso, como sabemos que f é contínua, teremos que, para um ε escolhido igual a $f(c) - u$, podemos sempre encontrar um δ tal que $|x - c| < \delta \implies |f(x) - f(c)| < \varepsilon = f(c) - u$. E isso significa que para qualquer valor de x no intervalo $(c-\delta,c+\delta)$ teremos que $f(x)$ vai estar mais perto de $f(c)$ que u, então certamente $f(x) > u$ para qualquer um desses elementos no intervalo $(c-\delta,c+\delta)$, e em particular para $c - \delta$. Mas se isso for verdade, então todo mundo entre $c - \delta$ e c não está em S e ninguém depois de c está em S (c é majorante de S), então $c - \delta$ é majorante de S e é menor que seu supremo c, o que é absurdo.

Se o caso fosse $f(c) < u$, teríamos uma situação parecida com a de cima, mas com a desigualdade trocada, teríamos que, para um ε escolhido igual a $u - f(c)$, podemos sempre encontrar um δ tal que $|x - c| < \delta \implies |f(x) - f(c)| < \varepsilon = u - f(c)$. E isso significa que para qualquer valor de x no intervalo $(c-\delta,c+\delta)$ teremos que $f(x)$ vai estar mais perto de $f(c)$ que u, então certamente $f(x) < u$ para qualquer um desses elementos no intervalo $(c-\delta,c+\delta)$, e em particular para $c + \delta$. Mas se isso for verdade, então todo mundo entre c e $c + \delta$ está em S, então c não é majorante de S, pois há elementos de S maiores que ele, o que é absurdo.

A única opção, portanto, é $f(c) = u$.

□

Mas é possível uma generalização dele? Sim, mas não é nada trivial. Você consegue provar que uma função contínua leva um conjunto conexo a um outro conjunto conexo, e que os únicos conexos dos reais são os intervalos, mas isso vai um pouco além do que eu pretendo contar hoje para você.

8 Conjuntos compactos

8.1 Definições

Para terminar esse texto, introduzo nosso último tipo de conjunto, que é de tremenda importância na análise: os conjuntos compactos. É bem difícil dar uma ideia intuitiva do que é um compacto, porque em espaços métricos bem gerais eles podem assumir uma forma estranha que escapa dessa noção intuitiva, mas dá para imaginar um compacto como um conjunto "pequeno" e cujas funções contínuas, neles, estão bem definidas, não explodem. Nós veremos que o intervalo [0,1] é compacto, enquanto o intervalo [0,1) não o é. Note que uma função contínua pode ir para onde quiser enquanto tende a 1, enquanto em [0,1] qualquer função contínua com esse domínio deve ser completamente "desenhável" em um gráfico.

A definição geral de compacto pode parecer estranha, mas é com ela que começo. Um conjunto compacto é aquele que pode ser recoberto por um número finito de conjuntos abertos. Essa ideia parece estranha, incluí uma figura para tentar explicar o que seria um recobrimento de abertos. O conjunto compacto, linha cheia, é recoberto por conjuntos abertos, que são as bolas amarelas, de forma a todos os pontos do compacto pertencerem a ao menos um aberto.

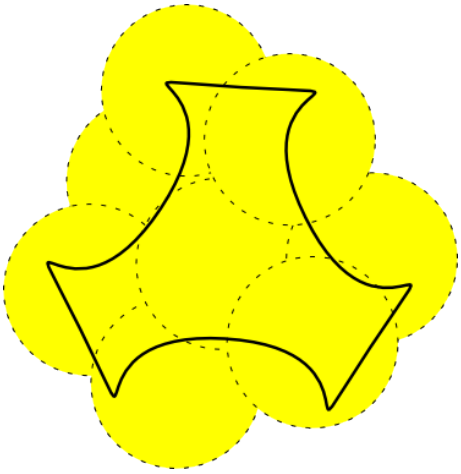


Figura 9: Representação de um recobrimento finito de abertos em um compacto.

Mas essa não é ainda a definição exata. Eu preciso garantir não somente que existe um recobrimento finito, mas que é impossível cobrir o espaço com uma cobertura de abertos tal que seja impossível usar apenas um número finito de elementos dessa cobertura para forrar o espaço todo. Ora, não costuma ser difícil encontrar um aberto único e imenso para cobrir o espaço todo, o que quero é provar que não existe nenhuma cobertura que precisa ser infinita para cobrir meu espaço todo. Se ela precisa ser infinita para funcionar, ele não é mais compacto.

Definição 18 (Recobrimento)**.** *Seja X um conjunto. Dizemos que a família de conjuntos $\{A_n\}$*

é um recobrimento, ou cobertura, de X se X estiver contido na união da família $\{A_n\}$, ou seja:

$$X \subset \bigcup_n A_n$$

Definição 19 (Conjunto compacto)**.** *Dizemos que um conjunto X é compacto se todo recobrimento seu composto de famílias de abertos possuir um subrecobrimento finito de abertos. Em outras palavras, seja X um conjunto e $\{A_n\}$ uma família de abertos, com $n \in I$ um subconjunto (finito, infinito ou até ele mesmo) dos naturais e*

$$X \subset \bigcup_{n \in I} A_n.$$

Se X é compacto, então existe um subconjunto $J \in I$ finito tal que:

$$X \subset \bigcup_{n \in J} A_n.$$

Eu confesso que essa não é a propriedade mais agradável do mundo de verificar, mas não se preocupe. Existem teoremas, em particular o de Heine-Borel, que nos facilitam bastante a vida, dizendo que em alguns espaços específicos os compactos ganham uma forma bem mais agradável e de fácil reconhecimento. Pelos teoremas que vou enunciar, você será capaz de perceber as propriedades lindas dos compactos, que os tornam conjuntos que são quase tudo que você sempre quis de um conjunto.

8.2 Resultados

Vamos começar facilitando nossa vida. Normalmente a ideia de compacto é inicialmente exposta na análise real, porque os subconjuntos compactos de $\mathbb{R}^n$ são bem fáceis de se encontrar, eles são os conjuntos fechados e limitados. É fácil ter uma imagem mental que um conjunto fechado e limitado: alguém cujas sequências não podem sair dele e cujas sequências não podem ir para o infinito, um conjunto "pequeno" e cujos elementos não tendem a ninguém fora dele. Cuidado, por mais que essa imagem mental seja agradável, ela em geral é falsa: a bola fechada de raio 1 centrada na origem, por exemplo, nunca é um compacto em dimensão infinita[3].

Mas antes de provar Heine-Borel, vamos precisar de um outro resultado, que também é importante, pois é também um bom jeito de achar compactos mais genéricos:

Teorema 17. *Todo subconjunto fechado de um compacto é compacto.*

[3]Esse é um resultado interessante, chamado teorema de Riesz, que é um jeito alternativo de definir o que é um espaço de dimensão infinita: alguém em que a bola unitária fechada não é compacta.

Demonstração. Vamos pensar na prova juntos. Eu tenho três espaços, $F \subset K \subset X$. X é o espaço todo, T é um compacto e F é um fechado, eu quero provar que F é compacto. Para isso, eu tenho que provar que, se eu pegar uma cobertura aberta de F, ela possui um subconjunto finito de seus recobrimentos que basta para forrar o espaço F.

Então vamos lá, seja $C_F = \{A_n\}$ uma cobertura de abertos de F, ou seja, cada A_n é um aberto e F está contido na união deles. Preciso agora provar que sempre acho um subconjunto de C_F com um número finito de elementos que cobre F. Seja $A = X - F$ o complementar de F, que é um aberto. Tomemos a família $C_T = C_F \cup \{A\} = \{A_n, A\}$, uma nova coleção de abertos. Mas dessa vez a união de todos eles cobre T (na verdade, cobre até X!), e são todos abertos. Como T é compacto, eu consigo extrair um número finito de elementos dessa família e formar uma $C_T{}'$ subfamília finita de elementos que recobre T e, de brinde, recobre F também (que está contido em T).

Como F não possui nenhum elemento de A, basta eu tirar o A dessa família $C_T{}'$ para ela se tornar uma família finita de abertos cuja união contém F, pois não tirei ninguém de F tirando A. Note que essa família terá apenas elementos que estavam originalmente em C_K, pois o que fiz foi: incluir A para achar uma cobertura de T, pescar alguns elementos que ainda recobrem T, tirar o A que coloquei. Isso é o equivalente a pescar um número finito de A_n's que cobrem F e são abertos, logo F é compacto. □

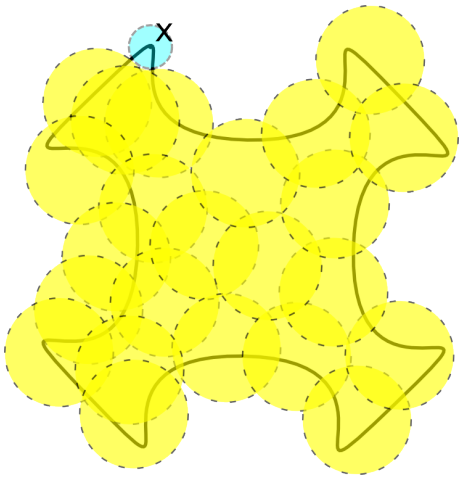


Figura 10: Representação da demonstração do fechamento de um compacto.

E eu vou precisar de mais dois resultados bem importantes:

Teorema 18. *Todo espaço métrico compacto é fechado.*

Demonstração. Seja K compacto, vamos provar que K é fechado. Para isso, a ideia é pegar um ponto de acumulação e provar que, se K é compacto, ele tem que estar em K. De fato, se K houvesse um ponto x que fosse seu ponto de acumulação e não estivesse nele, eu poderia escolher a cobertura: bolas abertas centradas em todos os $y \in K$ tais que, para algum $\varepsilon_y > 0$, eu posso definir $B_\delta(y) \cap B_\varepsilon(x) = \emptyset$. Ou seja, são as bolas abertas centradas em todos os pontos de K que são disjuntas de alguma vizinhança de x, note bem que essa vizinhança varia de ponto para ponto e é só por isso que a frase faz sentido. É evidente que essa família cobre K, mas é impossível extrair uma subfamília finita dela. Por quê?

Suponha que você tire um número finito dessas bolas abertas e tente cobrir K. Cada uma dessas bolas é disjunta de alguma vizinhança de x, mas, sendo um número finito, eu posso

extrair a menor das vizinhanças, que será disjunta de todas as bolas abertas da cobertura. Mas se x é ponto de acumulação de K, em toda bola aberta sua há ao menos um elemento de K, mas nenhuma cobertura pode "entrar" nessa bola, então nenhuma subcobertura finita é capaz de cobrir todo o K e K não seria compacto. Logo, a única opção é assumir que K contém todos os seus pontos de acumulação e, por isso, é fechado. □

Teorema 19. *Todo espaço métrico compacto é limitado.*

Demonstração. Essa não é uma demonstração difícil. Se K é compacto, basta pegar um recobrimento de bolas abertas centradas todas em um mesmo ponto $x \in K$ com raios tão grandes quanto eu preciso que eles sejam para cobrir o conjunto todo. Note que esse recobrimento existe e está bem definido, basta tomar como raios das bolas a distância de x ao ponto y, somar 1 nessa distância e pronto, terei uma bola aberta centrada em x que recobre y.

Como K é compacto, existe uma subcobertura finita de K contida nessa que defini, ou seja, um número finito dessas bolas abertas. Como ele é finito, eu pego a maior de todas, a de maior raio, e noto que todas as demais estão incluídas nela, ou seja, todo o K está dentro dela. Se o raio dela é R, teremos que $d(z,y) \leq d(x,y) + d(y,z) \leq D + D = 2D$, ou seja, existe um valor que majora todas as distâncias possíveis dos elementos de K, o que torna K limitado. □

Vamos usar esses resultados para provar que, no $\mathbb{R}^n$, os compactos assumem uma configuração bem mais agradável e fácil de verificar:

Teorema 20 (de Heine-Borel)**.** *Se $X \subset \mathbb{R}^n$, então X é compacto se, e somente se, X é fechado e limitado.*

Demonstração. Já provamos acima que, não apenas para $\mathbb{R}^n$, compacto implica fechado e limitado (os dois teoremas anteriores). Provemos agora que, em $\mathbb{R}^n$, fechado e limitado implica compacto. Seja K um conjunto fechado e limitado, contido em $\mathbb{R}^n$. Se ele é limitado, podemos colocar ele dentro de uma "caixa" cúbica da forma $T = [-a,a]^n$. Essa caixa, em $\mathbb{R}^2$, é um quadrado de lado $2a$, em $\mathbb{R}^3$ um cubo de aresta $2a$ e assim por diante.

Pelo teorema 17, todo subconjunto fechado de um compacto é compacto. Como $K \subset T$ e K é fechado, preciso apenas provar que a caixa é compacta. Para tal, vou argumentar por absurdo, vou fingir que a caixa não é compacta e gerar um absurdo com isso, deixando apenas como possibilidade a caixa ser compacta. Vamos, para isso, criar uma família de conjuntos da seguinte forma:

Se a caixa não é compacta, então existe uma cobertura de abertos C que não possui um recobrimento finito que seja suficiente para cobrir a caixa. Vou então cortá-la pela metade em todas as dimensões e obter 2^n caixas cujas dimensões valem metade da dimensão da caixa original, qualquer semelhança com Bolzano-Weierstrass não é mera coincidência. No caso de um quadrado, seria o equivalente a dividir em quatro quadrados iguais cujo lado é metade

do do quadrado original. No cubo, em oito cubos iguais cuja aresta é metade da do cubo original. Como eu preciso de um número infinito de abertos daquela cobertura para cobrir a caixa, em alguma dessas caixas menores eu ainda vou precisar de infinitos abertos para cobrir (do contrário, eu conseguiria cobrir a caixa toda com um número finito de abertos!). Seja a caixa original T_0, vamos pegar a caixa pequena que também exige infinitos e chamar de T_1.

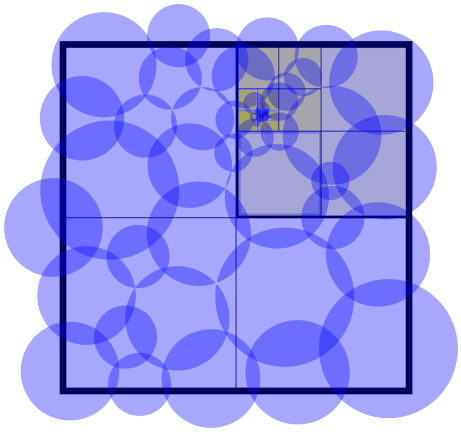


Figura 11: Representação de parte da demonstração de Heine-Borel.

Mais uma vez, eu pego T_1 e corto-a pela metade em todas as dimensões. Em T_0, a maior distância possível entre dois pontos era a diagonal da caixa n-dimensional, $a\sqrt{n}$. Em T_1, a maior distância entre dois pontos é $(a/2)\sqrt{n}$. Cortando mais uma vez, teremos alguma caixa que precisa de infinitos recobrimentos, chamaremos de T_2, seu diâmetro será $(a/4)\sqrt{n}$. Notemos que:

- Todas as caixas, pequenas ou grandes, são conjuntos fechados, completos, limitados e não-vazios.
- $T_{i+1} \subset T_i$, elas são encaixantes.
- $\text{diam}(T_i) = \frac{a}{2^i}\sqrt{n}$, que tende a 0 quando n cresce.

Ora, essas são exatamente as hipóteses do teorema da intersecção de Cantor (teorema 11). Por ele, sabemos que existe um p que pertence a todos os T_n's. Em particular, $p \in K$, então existe um aberto U da cobertura C que cobre p. Como U é aberto, p é ponto interior de U e podemos encontrar uma bola aberta centrada em p que está inteira contida em U, seja o raio dessa bola ε.

Ora, como o diâmetro dos T_i's tende a zero, uma hora ele ficará menor que ε, e ele conterá p, então $T_i \subset B_d(p) \subset U \in C$. Mas eu havia dito que T_n exigia um número infinito de abertos de C para ser recoberto, e eu acabo de provar que com apenas um, U, eu posso cobri-lo, o que é uma contradição. Dessa forma, a única possibilidade é a não-existência de uma cobertura desse tipo, toda cobertura deve possuir subcobertura finita e a caixa é compacta. Logo, K é compacto. □

O resultado seguinte já foi visto nos reais, aqui ele será apresentado com mais generalidade. O teorema de Bolzano-Weierstrass diz que, nos reais, toda sequência limitada possui subsequência convergente. Isso significava poder aprisionar toda a sequência em um conjunto limitado e, com cortes, ir fechando o limite para alguma subsequência dela. Mas esse resultado é mais geral, ele vale para todo espaço métrico compacto. Aqui, a compacidade nos dará as propriedades que, nos reais, tínhamos usando o fato de a sequência ser limitada e cortando a caixa de forma controlada. Em um espaço métrico geral, não podemos fazer cortes, mas podemos usar os recobrimentos para salvar o dia.

Essa propriedade, um conjunto em que toda sequência limitada sua possui subsequência convergente, recebe o nome de sequencialmente compacto. Nos espaços métricos, sequencialmente compacto e compacto são a mesma coisa, como vamos provar. Na verdade, vamos provar que quatro propriedades são a mesma coisa: ser compacto, ser sequencialmente compacto, qualquer conjunto infinito seu possuir um ponto de acumulação e ser completo e completamente limitado.

Definição 20 (Espaço totalmente limitado)**.** *Dizemos que um espaço métrico M é totalmente limitado quando, para todo $\forall \varepsilon > 0$, podemos decompor M como uma união finita de conjuntos X_i tais que todos os X_i possuem diâmetro menor que ε.*

Ou seja, um conjunto completamente limitado pode ser coberto por um número finito de conjuntos, por menores que esses conjuntos sejam. Um exemplo é o intervalo $[0,1] \subset \mathbb{R}$, que, para um dado $\varepsilon > 0$, tomamos $0 < \delta < \varepsilon$ e cobrimos o intervalo com $[n.\delta, (n+1)\delta]$. Em outras palavras, eu sempre consigo "quebrar" o conjunto em pequenos pedaços que, apesar de pequenos, ainda são finitos.

Com essa definição feita, podemos enunciar um importante teorema de caracterização dos compactos. Quando eu digo "são equivalentes as afirmações", digo que se uma é verdade sobre um conjunto, todas são. E a maneira de provar será interessante.

Teorema 21 (de Borel-Lebesgue)**.** *Seja K um espaço métrico. São equivalentes as afirmações:*

1. *K é compacto.*
2. *toda sequência em K possui subsequência convergente. (Bolzano-Weierstrass generalizado)*
3. *todo subconjunto de K que possui infinitos elementos contém um ponto de acumulação.*
4. *K é completo e totalmente limitado.*

Antes de provar esse teorema, note a força dele. Juntando aos outros, sabemos agora muitas coisas sobre os compactos. Todo compacto é fechado, limitado, completo, toda sequência

em si possui subsequência convergente e todo subconjunto infinito seu contém um ponto de acumulação, sabemos também que todo subconjunto fechado seu é compacto e que, no $\mathbb{R}^n$, basta que um conjunto seja fechado e limitado para que seja compacto. Provar que um conjunto é compacto nos dá, de brinde, diversas propriedades, além do teorema de Weiertrass que vou enunciar e provar daqui a pouco. Primeiro, provemos o teorema acima.

Demonstração. Tenho quatro propriedades, quero provar que são equivalentes. Não preciso provar seis "se, e somente se", o que seriam doze teoremas, basta eu efetuar uma "demonstração em cadeia". Vou provar que (1) $\implies$ (2) $\implies$ (3) $\implies$ (4) $\implies$ (1), o que liga todas as propriedades e as torna equivalentes.

(1) $\implies$ (2): Seja x_n uma sequência em K. Vamos supor, por absurdo, que ela não tem subsequência convergente. Então, para começar, o conjunto $\{x_n\}$ contém apenas pontos isolados, nenhum ponto de acumulação (se não, esse ponto seria o limite de uma subsequência). Assim, todo ponto de $\{x_n\}$ pode ser centro de uma bola aberta que contém apenas ele mesmo, e mais ninguém (do contrário, se sempre houvesse alguém na bola aberta, ele seria ponto de acumulação). O conjunto $F_0 = \{x_n\}$ é formado apenas de pontos isolados, então é fechado. Vamos usar essas bolas abertas com o complementar do F_0 para cobrir o K.

Temos várias bolas abertas e um aberto maior e mais gordo que consiste em todos os que não estão na sequência, a união disso certamente cobre K. Mas como K é compacto, deveria haver um subrecobrimento finito dele que consegue cobrir K todo. Mas note que isso é completamente impossível, a remoção de qualquer aberto dessa cobertura excluirá um ponto ao menos de S. A grande cobertura complementar de F_0 não pega nenhum elemento de S, cada bola aberta centrada em S pega apenas um elemento de S, não é possível extrair uma subcobertura finita e, por isso, K não seria compacto, logo não é possível que uma sequência possua apenas pontos isolados e, assim, ela deve possuir um ponto de acumulação, logo há ao menos uma subsequência sua com limite.

(2) $\implies$ (3): Sempre podemos, em um conjunto infinito, tomar uma sequência de elementos sempre diferentes entre si. Por (2), ela deve possuir subsequência convergente, que é um ponto de acumulação desse conjunto.

completar

□

Como último teorema envolvendo compactos, apresento uma propriedade fundamental que envolve compactos e continuidade, chamada de teorema de Weierstrass. Nele, vamos provar que toda função contínua definida em um compacto possui seu valor máximo e seu valor mínimo nesse compacto.

Teorema 22 (de Weierstrass)**.** *Toda função contínua real definida em um compacto assume, no compacto, seus valores de máximo e de mínimo*

Demonstração. Primeiro vamos provar que a imagem da função contínua definida em um compacto é limitada, depois vamos provar que o supremo e o ínfimo desse conjunto são imagem de algum elemento do domínio compacto.

Lema 2. *Seja K compacto e $f : K \to \mathbb{R}$ uma função contínua. Então a imagem de f em K, denotada $f(K)$, é um conjunto limitado.*

Demonstração. Se a imagem não fosse limitada, eu sempre poderia encontrar, para qualquer n natural, um $x_n \in K$ tal que $f(x_n) > n$. Eu vou definir com isso uma sequência, cada elemento x_n dela possui imagem maior que n. Sendo definida em K compacto, pela propriedade 2 do teorema 21, que é uma versão generalizada de Bolzano-Weierstrass, eu sempre posso encontrar uma subsequência x_{n_k} convergente. Sendo K compacto, ele também é fechado (teorema 17), então o limite dessa sequência estará em K, digamos então que $x_{n_k} \to x \in K$.

Mas já provamos que funções contínuas levam sequências convergentes em sequências convergentes (teorema 7), então se $x_{n_k} \to x$, teremos que $f(x_{n_k}) \to f(x)$. Mas eu e você sabemos que $f(x_{n_k}) > n_k \geq k$, por definição, essa sequência não converge de jeito nenhum, ela explode e fica tão grande quanto se queira. Logo, a única opção é a imagem ser limitada. □

Tendo provado que $f(K)$ é limitado, sabemos que, sendo um subconjunto dos reais, possui supremo e ínfimo. Vamos agora provar que há elementos em K que são levados tanto no supremo quanto no ínfimo. Seja M o supremo de $f(K)$. Então, para todo N natural, existem elementos no domínio enviados entre M e $M - 1/N$ (se assim não fosse, $M - 1/N$ seria o supremo de $f(K)$). Podemos pegar, para cara N, um elemento que é enviado nesse intervalo, definindo com isso usa sequência x_n tal que $M - 1/n \leq f(x_n) \leq M$.

Essa sequência x_n está toda em K, então, pela propriedade 2 do teorema 21, ela possui uma subsequência convergente $x_{n_k} \to x$. Como K é compacto, é fechado (teorema 17) e $x \in K$. Sendo f contínua, ela leva sequências convergentes em sequências convergentes (teorema 7), então $f(x_{n_k}) \to f(x)$. Mas sabemos que $M - 1/n \leq f(x_{n_k}) \leq M$, os limites das funções das pontas é M, então, pelo teorema do confronto (teorema 1), $f(x_{n_k}) \to M$. Pela unicidade do limite (teorema 2), $f(x) = M$ e já provamos, porque K é compacto então fechado, que $x \in K$ (teorema 3). □

9 Normas

9.1 Definições

Ainda há algumas coisas a se falar sobre distâncias e sobre espaços métricos, mas é oportuno o momento. Na motivação eu também citei o problema de medir o tamanho de determinados

entes matemáticos, como matrizes ou funções, e nós ainda não lidamos com isso. Sabemos calcular a distância entre dois elementos, mas a ideia de tamanho é mais sutil. Sabendo distância, é possível definir tamanho?

A resposta é não. Para uma pessoa, por exemplo, o tamanho significa a distância em relação ao chão. Nos espaços métricos, não é clara a ideia de chão, não possuímos um elemento privilegiado que podemos adotar como referência e, ainda que pudéssemos, teríamos outros problemas. Duas pessoas iguais juntas medem o dobro de uma pessoa sozinha, mas não podemos aplicar essa ideia às distâncias, pois não podemos "somar" dois elementos de um conjunto arbitrário. Essa ideia não nos é estranha, não é claro como poderíamos somar pontos em $\mathbb{R}^2$, um ponto mais outro ponto não é uma operação que faz muito sentido.

Para isso, precisamos pensar em outro tipo de conjuntos. Quero conjunto que, ainda que eu não consiga pensar em distância, posso somar elementos. Não somente isso, eu preciso daquela ideia de "dobro", eu preciso poder multiplicar os elementos desse conjunto por números. Que números? Reais? Racionais? Para responder a essa pergunta, preciso também definir quem são esses números a serem multiplicados pelos elementos dos conjuntos.

Os itens desse capítulo não serão exatamente abordados em um curso de análise, que é o tema principal desse texto, mas no curso de álgebra linear. Ainda, o conceito de norma é fundamental para qualquer coisa que você pensar em fazer na matemática, por isso ela é de apresentação necessária.

Definição 21 (Corpo)**.** *Seja X um conjunto. Dizemos que X é um corpo se ele for munido de duas operações, que denotaremos $+$ e $*$, tais que:*

1. *X é fechado em $+$ e $*$: ou seja, para todo a,$b \in X$, $a + b \in X$ e $a * b \in X$.*
2. *$+$ e $*$ são associativas: $(a + b) + c = a + (b + c)$ e $(a * b) * c = a * (b * c)$.*
3. *$+$ e $*$ são comutativas: $a + b = b + a$ e $a * b = b * a$.*
4. *$+$ e $*$ possuem elemento neutro, ou seja, alguém que, $\forall x \in X$, $x * e_* = e_* * x = x$ e um outro alguém tal que $x + e_+ = e_+ + x = x$.*
5. *$+$ e $*$ são sempre inversíveis: para todo $x \in X$, sempre existe um $z \in X$ tal que $x+z = e_+$. Para qualquer outro x que não seja e_+, sempre existe um $y \in X$ tal que $x * y = e_*$.*
6. *A propriedade distributiva: $a * (b + c) = a * b + a * c$ é verificada.*

Para que as propriedades acima fiquem mais claras, costumamos chamar e_+ de 0, e_* de 1, o inverso de x em $+$ chamamos de $-x$ e o inverso de x em $*$ chamamos de x^{-1}. Eu reforço aqui que as operações $+$ e $*$, apesar de receberem muitas vezes o nome de soma e produto,

não precisam necessariamente ser a soma e o produto a que você está acostumado. Se elas satisfizerem as propriedades acima, podem ser qualquer coisa.

Não sei até que ponto eu realmente deveria ter introduzido isso, porque é uma definição meio abstrata, mas é importante você perceber que nós não precisamos sempre pegar os números reais para multiplicar os elementos do conjunto, precisamos apenas de um cara que satisfaça essas propriedades. Quando eu digo a palavra corpo, tenha em mente os reais, eles são o melhor exemplo de corpo com as operações de $+$ e $*$ convencionais. Mas eles não são o único.

Os racionais, com as operações convencionais de soma e produto, também são um corpo. Os inteiros, por outro lado, não são: falta-lhes o inverso multiplicativo. Você nunca consegue multiplicar 2 por um inteiro de forma a torná-lo 1, que é o neutro multiplicativo. As matrizes quadradas inversíveis $n \times n$ também não são um corpo, pois apesar de possuírem inverso elas não são comutativas no produto.

Um caso legal de corpo é o conjunto $\mathbb{Z}/p\mathbb{Z}$ com p um número primo. Para entender o que é um conjunto desses, dou um exemplo, o $\mathbb{Z}/5\mathbb{Z}$. Esse conjunto é composto dos "restos da divisão de um número por 5", ou seja, ele é $\{0, 1, 2, 3, 4\}$ e você opera normalmente esses caras, com soma e produto, mas há a regra: toda vez que você ultrapassa o 5, subtrai 5. Qual é a ideia? Se você soma um número cujo resto da divisão por 5 é 2 com um cujo resto da divisão por 5 é 3, você terá um cujo resto da divisão por 5 é 0, ele será divisível por 5. Por isso, nesse conjunto, $2 + 3 = 0$ e 3 é o inverso de 2. Não provo, mas enuncio, que esse tipo de conjunto só é corpo quando p é primo, esse é um resultado interessante da álgebra.

Tendo introduzido isso, digo que, toda vez que você ler "um corpo K", substitua isso na sua cabeça pelos números reais, isso dará conta de 80% dos casos. Os outros 20% são quando estamos nos divertindo com conjuntos interagindo com os complexos, o que é a base da mecânica quântica, por exemplo.

Em seguida, vou definir um tipo especial de conjunto. São elementos que podem ser somados entre si (como matrizes, vetores, funções) e que podem ser multiplicados por um número (como matrizes, vetores e funções também). Note que não é tão evidente a noção de multiplicar um elemento por outro (o que é um vetor vezes outro vetor?), mas faz bastante sentido multiplicá-lo por um "número", o equivalente a dobrá-lo de tamanho, sendo esse número um elemento de algum corpo que eu escolher (pensou nos reais? Pois deveria!).

Definição 22 (Espaço vetorial)**.** *Sejam V um conjunto e K um corpo. Dizemos que V é um K-espaço vetorial, e chamamos seus elementos vetores, se podemos nele definir uma operação $+$ que:*

1. *seja associativa: $u + (v + w) = (u + v) + w$, $\forall u, v, w \in V$.*
2. *seja comutativa: $u + v = v + u$, $\forall u, v \in V$.*
3. *possua um elemento neutro, denotado vetor nulo: $u + 0 = o + u = u$, $\forall u \in V$.*

4. seja inversível: para todo u sempre existe um $-u$ tal que $u + (-u) = 0$.

e também uma operação entre V e K, chamada multiplicação por escalar, tal que:

5. $\lambda(\mu.v) = (\lambda\mu)v$, $\forall\lambda, \mu \in K$ e $\forall v \in V$.

6. $\lambda(u + v) = \lambda.u + \lambda.v$, $\forall\lambda \in K$ e $\forall u, v \in V$.

Essa grande quantidade de definições e propriedades não deve te assustar. É natural que eu precise disso: não é nada evidente a ideia de poder somar coisas, eu preciso definir com carinho quando quero deixar essas "coisas" o mais geral possível.

Muita coisa é um espaço vetorial. Essencialmente, basta que eu possa somar elementos e multiplicá-los por números para ter um espaço vetorial: matrizes, funções, soluções de uma equação diferencial homogênea, estados de um sistema quântico, vetores tangentes a uma variedade em um ponto, os reais, os complexos, é uma noção bem geral e todas as conclusões deduzidas dessas propriedades acima se aplicam a qualquer um desses entes matemáticos que entram no critério de espaço vetorial, ainda que você saiba o que é metade deles.

Um espaço vetorial possui o que chamamos de "dimensão". Não quero entrar em detalhes, isso é assunto para álgebra linear e não para mim, mas vou precisar definir isso para mostrar uma propriedade importante sobre normas. A dimensão de um espaço vetorial é o número mínimo de vetores de que você precisa para formar com somas e múltiplos deles o espaço todo. O $\mathbb{R}^2$, por exemplo, pode ter cada membro seu construído apenas com múltiplos e somas dos vetores $\binom{1}{0}$ e $\binom{0}{1}$. Ora, é fácil ver que $\binom{a}{b} = a\binom{1}{0} + b\binom{0}{1}$. O $\mathbb{R}^2$ possui, por isso, dimensão dois.

As funções contínuas, por outro lado, possuem dimensão infinita: você nunca consegue um número finito de funções tais que, apenas com multiplicações por reais e somas dessas funções, você gere todo o espaço de funções contínuas.

Tendo definido esse tipo de objeto, posso já pensar em definir uma função que meça o tamanho deles.

Definição 23 (Norma)**.** *Seja V um K-espaço vetorial, com K um corpo contido nos complexos. Diz-se norma uma função $\|\cdot\| : V \to \mathbb{R}$ que satisfaz as propriedades:*

1. $\|\lambda v\| = |\lambda|\,\|v\|$, $\forall\lambda \in K$ e $\forall v \in V$.

2. $\|u + v\| \leq \|u\| + \|v\|$, ou seja, ela obedece à desigualdade triangular.

3. $\|v\| = 0 \Leftrightarrow v = 0$, ou seja, ela é definida.

Para aproveitar todos os demais resultados do texto, vamos provar que normas e distâncias possuem uma relação direta.

Teorema 23. *Toda métrica pode ser usada para definir uma distância da forma* $d(u,v) = \|u - v\|$. *Logo, todo espaço vetorial normado é métrico.*

Demonstração. Basta provar que $\|x - y\|$ satisfaz as propriedades de uma distância. Vejamos, vamos primeiro provar que $d(x,y) = 0 \Leftrightarrow x = y$. A recíproca é evidente, pois $\|x - x\| = \left\|\vec{0}\right\| = 0.\left\|\vec{0}\right\| = 0$. Estou denotando o vetor nulo com essa flecha apenas para diferenciá-lo do neutro da adição do corpo K. Os neutros de espaços diferentes são coisas diferentes, oras. Falta provar que se $d(x,y) = 0$, então $x = y$. É certo que $d(x,y) = \|x - y\| = 0$, o que implica que $x - y = 0$ pela propriedade 3 da norma. Para argumentar que $x = y$, vamos ter que provar um pequeno lema:

Lema 3. *O inverso do elemento* x *é único, denotado* $-x$.

Demonstração. Suponha que um elemento x possua dois inversos, $-x$ e y. Então $x - x = 0$ e $x + y = 0$. Pegando essa segunda equação e somando $(-x)$ dos dois lados, o que preserva a igualdade, teremos $-x + (x + y) = -x$. Pela associatividade, podemos escrever $-x + (x + y) = (-x + x) + y = y = -x$, o que prova que o inverso é único. □

Sendo o inverso único, podemos concluir que $\|x - y\| = 0 \Leftrightarrow x = y$. A segunda propriedade é facilmente satisfeita pela comutatividade da soma do espaço vetorial: $d(x,y) = \|x - y\| = \|(-1)(y - x)\| = \|y - x\| = d(y,x)$. Para a desigualdade triangular, partimos da propriedade equivalente nas normas e somamos zero:

$$d(u,v) = \|u - v\| = \|u - w + w - v\| \leq \|u - w\| + \|v - w\| = d(u,w) + d(v,w)$$

□

E, com isso, podemos importar todos os resultados anteriores, a noção de aberto, fechado, completo e de função contínua para espaços vetoriais normados.

É importante citar um tipo de espaço recorrente na matemática. Todo espaço vetorial normado completo, que é quase tudo o que você sempre quis de um espaço, recebe um nome especial.

Definição 24 (Espaço de Banach)**.** *Dizemos um espaço de Banach todo espaço vetorial normado completo.*

Juntando todos esses conceitos, podemos fazer algumas perguntas pertinentes. Você sabe que, dependendo da distância, abertos em uma podem não ser abertos na outra. Vou dar um exemplo. Vou usar uma distância cruel, a chamada métrica discreta, a que associa o número 1

a dois pontos diferentes e zero à distância de pontos iguais, ou seja:

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

Nessa métrica, ninguém é aberto pois qualquer bola de raio $r = 1/2$ centrada em um ponto contém apenas o ponto, então nenhum ponto é ponto de interior. Para métricas, no entanto, as coisas mudam um pouco. A noção de aberto e fechado ainda depende da métrica, mas as coisas ficam um pouco diferentes. Vale a penas estudar quais métricas geram os mesmo abertos e quais não, quais possuem a mesma noção de convergência ou não, o que inspira a definição:

Definição 25 (Normas equivalentes)**.** *Duas normas $\|\cdot\|_1$ e $\|\cdot\|_2$ de um espaço vetorial V são ditas equivalentes se existes constantes C e D tais que, para todo $x \in V$:*

$$C \|x\|_1 \leq \|x\|_2 \leq D \|x\|_1$$

Repare que se uma delas converge para zero quando x vai para algum lugar, a outra também vai, elas estão intimamente ligadas. Como nós vamos provar, isso define uma equivalência nos abertos também.

9.2 Resultados

Vamos começar tratando de métricas equivalentes. Um teorema fundamental, que facilita bastante nossa vida em espaços de dimensão finita, é o que demonstra a equivalência de todas as métricas em dimensão finita. Sua utilidade é evidente, para demonstrar qualquer propriedade de espaços em dimensão finita, podemos pegar uma norma qualquer, provar para ela e teremos provado para todas.

Então alguns teoremas são necessários, porque esse não é um resultado fácil e exige algum esforço, vamos fazer por partes. Primeiro, vamos provar que se duas normas são equivalentes, os abertos em uma são abertos em outra. Depois, vamos mostrar que se uma norma é equivalente a outra e essa outra é equivalente a uma terceira, então a primeira é equivalente à terceira. Por fim, vamos provar que, em dimensão finita, todas as normas são equivalentes.

Teorema 24. *Se duas normas são equivalentes, os abertos em uma são abertos em outra.*

Demonstração. Sejam $\|\cdot\|_1$ e $\|\cdot\|_2$ duas normas definidas em V espaço vetorial. As distâncias decorrentes delas são d_1 e d_2. Seja $A \subset V$ um aberto de V de acordo com a norma $\|\cdot\|_1$, teremos então que, dado $x \in A$, existe uma bola aberta $B^1_\varepsilon(x) \subset A$, notei com o índice ^{1} a métrica associada à bola. Isso significa que $a \in B^1_\varepsilon(x) \implies \|x - a\|_1 < \varepsilon$.

Mas sabemos que existe um D real tal que $\|x-a\|_1 \leq D\|x-a\|_2$, então $\|x-a\|_2 < \frac{\varepsilon}{D}$. Assim, podemos definir uma bola centrada em x na métrica $\|\cdot\|_2$ com raio $\frac{\varepsilon}{D}$ que está inteira em A, logo A é um aberto na métrica $\|\cdot\|_2$. Se começássemos com um aberto em $\|\cdot\|_2$, bastaria definir uma bola de raio $\frac{\varepsilon}{C}$ na norma $\|\cdot\|_1$ para provar que A também é um aberto em $\|\cdot\|_1$. □

Pronto, provamos o teorema que justifica a que as normas equivalentes vieram. Distâncias definidas através dessas normas produzem os mesmos abertos, o que nos garante uma certa estabilidade ao considerarmos conjuntos nesses espaços. O teorema que segue, a prova que de em dimensão finita todas as normas são equivalentes, amplifica bastante o poder desse teorema anterior, mas sua demonstração não é simples. Ela será feita da seguinte forma: eu vou encontrar uma métrica privilegiada nos espaços vetoriais de dimensão finita e provar que ela é equivalente a qualquer outra métrica.

Mas, para provar que duas são equivalentes entre si, não basta provar que são equivalentes a uma terceira se eu não provar o teorema que segue:

Teorema 25. *Dadas três normas $\|\cdot\|_1$, $\|\cdot\|_2$ e $\|\cdot\|_3$ definidas em um espaço vetorial. Se $\|\cdot\|_1$ é equivalente a $\|\cdot\|_2$ e $\|\cdot\|_2$ é equivalente a $\|\cdot\|_3$, então $\|\cdot\|_1$ é equivalente a $\|\cdot\|_3$.*

Demonstração. Se $A\|x\|_1 \leq \|x\|_2 \leq B\|x\|_1$ e $C\|x\|_2 \leq \|x\|_3 \leq D\|x\|_2$, então $AC\|x\|_1 \leq \|x\|_3 \leq BD\|x\|_1$. □

Finalmente, o teorema principal dessa seção.

Teorema 26. *Em dimensão finita, todas as normas são equivalentes.*

Demonstração. Como prometi, vamos caçar uma métrica bem simpática que tenha a propriedade de ser compatível com todo mundo. Se estamos tratando de um espaço de dimensão finita, podemos sempre escolher uma base nele. Quando falamos de dimensão, comentamos que era "o número mínimo de vetores de que você precisa para formar, com somas e múltiplos deles, o espaço todo". Uma base do espaço vetorial é um desses conjuntos capaz de formar o espaço todo com o menor número de elementos possíveis.

E cada elemento seu pode ser escrito como a soma desses elementos da base vezes alguns coeficientes seus, ou seja, se o conjunto $\{a_n\}$ é base para V, então todo $x \in V$ pode ser escrito como $x = x_1a_1 + x_2a_2 + \ldots + x_na_n$, onde x_n são elementos do corpo, ou seja, números reais na maior parte dos casos.

Escolhida essa base, a norma privilegiada que vamos escolher é: $\|x\|_\star = \sqrt{\sum_n x_n^2}$. Nos reais, ela é a norma euclidiana tradicional, a que já estamos tão acostumados. Mas ela pode ser definida em qualquer espaço de dimensão finita, já que todos eles podem ser colocados em uma base e essa norma usa seus coeficientes. De fato, ela depende da base, mas vamos provar que ela é equivalente a todas as outras para qualquer base escolhida.

E $\|x\|_\star$ é de fato uma norma. É linear, positiva (dificilmente uma raiz de soma de quadrados será negativa) e zero apenas quando todos os componentes do vetor forem zero (ou seja, o vetor é o vetor nulo). A desigualdade triangular é levemente mais complicada, mas uma pequena explorada em desigualdades com a raiz quadrada dá conta do recado.

Assim, seja $\|\cdot\|_1$ uma norma em V um espaço vetorial, provemos que $\|\cdot\|_1$ é equivalente a $\|\cdot\|_\star$.

Primeira parte: provar que existe M tal que $\|x\|_1 \leq M \|x\|_\star$.

Seja $x \in V$. Como estamos em dimensão finita, podemos escrever x como seus componentes ndaquela base $\{a_n\}$: $x = \sum_n x_n a_n$. Assim:

$$\|x\|_1 = \left\| \sum_n x_n a_n \right\| \overset{1}{\leq} \sum_n |x_n| \|a_n\|_1 \overset{2}{\leq} \sqrt{\sum_n |x_n|^2} \sqrt{\sum_n \|a_n\|_1^2} = \sqrt{\sum_n \|a_n\|_1^2} \|x\|_\star$$

Vale a pena comentar as passagens 1 e 2. Na 1, usei um combinado de linearidade da norma com desigualdade triangular para separar os elementos da base de seus coeficientes e de si próprios. Na 2, usei uma relação aritmética um pouco complicada, mas que é válida e é até demonstrada nesse texto um pouco mais para frente, é, aliás, o teorema 27. Sei que não é legal usar o teorema 27 para provar o 26, mas garanto que o primeiro não precisa do último para funcionar, então estou autorizado.

O valor de $\sqrt{\sum_n \|a_n\|_1^2}$ não depende de x, apenas da base escolhida e da norma $\|\cdot\|_1$. Como é uma soma finita de elementos finitos, é finita e tem algum valor, chamaremos de M. Note que aquela raiz dos x_1 é nossa norma favorita escolhida anteriormente. Assim, provamos que $\|x\|_1 \leq M \|x\|_\star$.

Segunda parte: provar que existe m tal que $m \|x\|_\star \leq \|x\|_1$.

Essa demonstração é mais sutil. Seja n a dimensão de V., primeiro vamos definir um conjunto S de $\mathbb{R}^n$ tal que:

$$S = \left\{ (a_1, \ldots, a_n) \in \mathbb{R}^n \mid \sum_j a_j^2 = 1 \right\}$$

Tal conjunto é um fechado em $\mathbb{R}$, como você pode provar usando uma técnica bem parecida à provar que o conjunto das matrizes com determinante nulo é um fechado em $M_n(\mathbb{R})$, como já fizemos. E parece bem fácil ver que esse conjunto é limitado, porque a norma de qualquer elemento dele não só não é infinita como é 1.

□

10 Produto interno

10.1 Definições

O último tema de que pretendo falar é a noção de produto interno. Ele é bem importante para praticamente todas as áreas da matemática e da física, pois está intimamente associado às noções de projeção, ângulo e decomposição de um elemento em outros mais simples. Muitas vezes, algumas estruturas matemáticas são complicadas demais para serem tratadas como são e escrevê-las como somas de elementos mais simples pode ajudar bastante.

Não à toa, esse tópico é mais próximo da álgebra linear, que não é exatamente o conteúdo desse texto. Álgebra linear lida com operações em espaços vetoriais, em que a noção de soma existe, e linearidade, que é uma generalização da "distributiva", passa a ser protagonista. Não vamos falar só de álgebra linear, mas em como ela pode nos ajudar a lidar com funções, conjuntos e outros elementos que já discutimos aqui.

Para falar de produto interno, precisamos já de um espaço vetorial. Tendo, vamos definir nele uma função da seguinte forma:

Definição 26 (Produto interno)**.** *Seja V um espaço vetorial sobre $\mathbb{R}$ ou $\mathbb{C}$. Dizemos que uma função $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ ou $\mathbb{C}$ é um produto interno sobre V se ela obedece às propriedades:*

1. *Simétrico conjugado: $\langle x, y \rangle = \overline{\langle y, x \rangle}$. Note que, para $\mathbb{R}$, isso significa que o produto interno é simétrico.*
2. *Linearidade no primeiro argumento: $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$ e $\langle x + y, z \rangle = \langle x, y \rangle + \langle y, z \rangle$.*
3. *Positivo e definido: $\langle x, x \rangle \geq 0$ e apenas igual a zero se $x = 0$, o elemento neutro do espaço vetorial.*

Em primeiro lugar, quero fazer você perceber que essas propriedades são bem mais fáceis de verificar que as propriedades de norma ou distância, basta achar uma função de duas variáveis linear na primeira, simétrica (ou quase simétrica) e que leva duas variáveis iguais em um número positivo. Não precisa mais checar, por exemplo, a desigualdade triangular.

Vale também notar que as duas primeiras propriedades implicam no que chamamos de "anti-linearidade" do segundo argumento: $\langle x, \lambda y \rangle = \overline{\langle \lambda y, x \rangle} = \bar{\lambda} \overline{\langle y, x \rangle} = \bar{\lambda} \langle x, y \rangle$. Podemos "tirar" um real ou complexo multiplicando o segundo elemento como o conjugado complexo dele.

Diferentemente da norma, o produto interno leva dois elementos do espaço vetorial a um elemento do corpo do espaço vetorial. Não faz muito sentido usar corpos diferentes ou extravagantes para definir produtos internos complicados, vamos ficar apenas com reais a complexos.

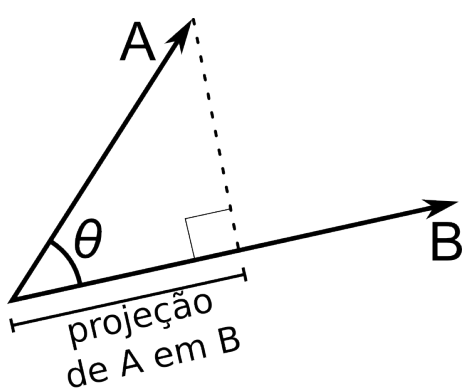


Figura 12: Representação da proje-

Lembro que a norma sempre leva elementos aos reais positivos, o produto interno pode levar a qualquer elemento do corpo, ainda que sempre leve o produto interno de um elemento consigo mesmo a um número positivo.

A razão dessa imposição é fácil de entender quando percebemos a razão da existência do produto interno. Ele sintetiza a ideia de projeção, de saber "o quanto de um elemento está em outro" e vice-e-versa. Essa ideia é bem natural quando falamos de vetores em $\mathbb{R}^2$, como você já deve ter visto. É fácil ver o quanto de um vetor você tem no outro, essa é a noção intuitiva de projeção ortogonal que temos, a "sombra" de um vetor na direção do outro. O produto interno, como nós veremos, nos dá a medida dessa projeção. No $\mathbb{R}^3$ ele nos dá uma fórmula simples e bonita, para espaços mais complicados não é tão trivial, mas a ideia de projeção ainda existe.

Espaços com produto interno são extremamente importantes. Tanto assim o é, damos um nome especial a espaços completos e com produto interno, são os espaços de Hilbert. Antes que você levante a mão para reclamar na falta da definição de distância, antecipo que todo espaço que possui produto interno possui também norma e, claro, distância, podemos usar o produto interno para definir todas essas coisas e, com isso, ganhar o direito de usar todo o resto do texto nesses espaços tão especiais.

Definição 27 (Espaço de Hilbert)**.** *Seja V um espaço vetorial sobre $\mathbb{R}$ (ou $\mathbb{C}$). Se V possui produto interno e é completo, então V é um espaço de Hilbert.*

10.2 Resultados

Eu sei que já deveria ter enunciado e provado que todo produto interno permite a definição, consigo, de uma norma e uma distância, mas para o fazer preciso de um resultado fundamental, que talvez seja uma das desigualdades mais importantes da matemática. Falo, claro, da desigualdade de Cauchy-Schwarz.

Teorema 27 (Desigualdade de Cauchy-Schwarz)**.** *Dado um produto interno $\langle \cdot, \cdot \rangle$ sobre um espaço V. Então, para quaisquer x,$y \in V$, vale a desigualdade:*

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle.$$

Demonstração. Se $y = 0$, é fácil ver que isso é verdade (os dois lados dão zero). Suponho então que $y \neq 0$, ou seja, $\langle y, y \rangle \neq 0$. Vou então definir um número $\lambda = \frac{\langle x, y \rangle}{\langle y, y \rangle}$. Podemos nos divertir, então, em calcular o produto interno $\langle x - \lambda y, x - \lambda y \rangle$. Não sabemos nada sobre tal produto,

mas uma coisa é certa, isso deve ser algo positivo, pela terceira propriedade do produto interno. Assim:

$$0 \leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - \lambda \langle y, x \rangle - \bar{\lambda} \langle x, y \rangle + |\lambda|^2 \langle y, y \rangle$$

Note que os dois últimos elementos dessa expressão se anulam quando você substitui λ por seu valor verdadeiro: $-\bar{\lambda}\langle x, y \rangle + |\lambda|^2 \langle y, y \rangle = \bar{\lambda} \left(-\langle x, y \rangle + \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle y, y \rangle \right) = 0$. Ficaremos apenas com os dois primeiros, em que podemos substituir o valor de λ:

$$0 \leq \langle x, x \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle y, x \rangle \implies |\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle$$

□

Para que isso faça sentido com o resto do texto, vou provar um resultado importante, parecido com aquele que diz que de toda norma um gero uma distância.

Teorema 28. *Todo produto interno define uma norma da forma* $\|x\| = \sqrt{\langle x, x \rangle}$*. Ou seja, todo espaço vetorial com produto interno é normado e é um espaço métrico.*

Demonstração. Para isso, vamos provar que uma norma definida daquela maneira é de fato uma norma, vamos provar as três propriedades delas a partir das propriedades do produto interno.

1. Basta notar que $\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\lambda \bar{\lambda} \langle x, x \rangle} = |\lambda| \, \|x\|$.

2. Ao invés de provar $\|x + y\| \leq \|x\| + \|y\|$, vou provar que $\|x + y\|^2 \leq (\|x\| + \|y\|)^2$, como todo mundo é positivo essas duas afirmações são equivalentes, quero só me livrar do trabalho de escrever muitas raízes quadradas. $\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle$. Esses dois elementos do meio são conjugados um do outro, a soma deles é o dobro da parte real deles (a parte imaginária se cancela, pois muda de sinal).

   E é verdade que um número é sempre menor ou igual a seu módulo, ainda mais se somarmos alguém ao módulo, ou seja: $x \leq \sqrt{x^2 + y^2}$. O que quero dizer é que $2\Re(\langle x, y \rangle) \leq 2|\langle x, y \rangle|$, a parte real é sempre menor que o módulo do número complexo. Assim, teremos que $\|x + y\|^2 \leq \langle x, x \rangle + 2|\langle x, y \rangle| + \langle y, y \rangle$. Mas esse termo do meio, *pela desigualdade de Cauchy-Schawrz* (teorema 27), é menor que o produto dos produtos internos, ou seja: $\|x + y\|^2 \leq \langle x, x \rangle + 2|\langle x, y \rangle| + \langle y, y \rangle \leq \langle x, x \rangle + 2\langle x, x \rangle \langle y, y \rangle + \langle y, y \rangle = \left(\sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle} \right)^2 = (\|x\| + \|y\|)^2$.

3. Decorre trivialmente da terceira propriedade do produto interno.

□

10.3 Exemplos de produto interno e norma

Vamos mostrar agora algumas aplicações e um pouco do poder de toda essa teoria. Esse é um prelúdio ao próximo capítulo, que lista alguns exemplos de para que tudo isso serve. Comecemos por apresentar alguns espaços de alguns produtos internos para esses espaços. Comecemos pelos mais simples, o espaço $\mathbb{R}^2$ e o $\mathbb{R}^3$. O produto interno natural desses espaços é conhecido como "produto escalar": $\langle a, b\rangle = a_1b_1 + a_2b_2$ para o $\mathbb{R}^2$ e com um termo extra para o $\mathbb{R}^3$, sendo a_1 e a_2 os componentes do vetor na sua base favorita (acredite se quiser, o produto interno não depende da base, mas essa demonstração fica para a álgebra linear). É fácil verificar as propriedades de produto interno para essa função. Ela gera a norma euclidiana convencional $\|a\| = \sqrt{a_1^2 + a_2^2}$ e essa norma gera a distância convencional no $\mathbb{R}^2$, aquela do teorema de pitágoras. O $\mathbb{R}^2$ é completo nessa norma, então o é em todas as normas, pois possui dimensão finita. Assim, o $\mathbb{R}^2$ (na verdade, o $\mathbb{R}^n$) é um espaço de Hilbert.

O espaço das matrizes reais quadradas $n \times n$: $M_n(\mathbb{R})$, munido do produto interno $\langle A, B\rangle = \operatorname{tr}(AB^t)$. Esse produto interno é conhecido como "produto interno de Frobenius". O traço de uma matriz, que é a soma de seus elementos da diagonal principal, é uma função linear, então as duas primeiras propriedades do produto interno já estão verificadas de graça para você. A terceira que $\operatorname{tr}(AA^t) \geq 0$, não é tão simples de verificar, mas confie em mim nessa, ela é verdade. Não vou ficar arrancando meus cabelos para te provar isso, quero apenas exemplos interessantes.

Ainda que você não esteja familiarizado com matrizes e traços de matrizes, veja o que você já sabe sobre elas. Sabe que $|\operatorname{tr}(AB^t)|^2 \leq \operatorname{tr}(AA^t)\operatorname{tr}(BB^t)$ (Cauchy-Schwarz), sabe que $\sqrt{\operatorname{tr}(AA^t)}$ define uma norma e que $\sqrt{\operatorname{tr}((A-B)(A-B)^t)}$ define uma distância, que isso pode ser usado para comparar duas matrizes e que podemos pensar em convergências e funções contínuas usando essa distância. Como esse espaço é de dimensão finita, todas as normas nele são equivalentes e qualquer aberto nele é aberto em qualquer outra norma. Teremos também que $GL_n(\mathbb{R})$ é um aberto nele, como provamos, isso significa que, para qualquer matriz inversível, podemos encontrar uma também inversível que esteja tão perto dela quanto quisermos, ou seja, que, para um ε escolhido, encontramos B inversível tal que $\sqrt{\operatorname{tr}((A-B)(A-B)^t)} < \varepsilon$. Também sabemos, pela desigualdade triangular, que, qualquer que seja a matriz C,

$$\sqrt{\operatorname{tr}((A-B)(A-B)^t)} \leq \sqrt{\operatorname{tr}((A-C)(A-C)^t)} + \sqrt{\operatorname{tr}((B-C)(B-C)^t)}$$

Deu para perceber que ganhamos diversas informações sobre as matrizes apenas provando que a operação é linear e positiva, esse é o poder da teoria que desenvolvemos. Provar qualquer uma dessas relações para o caso específico do traço poderia tomar tempo e certamente exigiria alguma habilidade, mas não precisamos disso, sabemos que elas valem para qualquer produto interno, que gera norma, que gera distância.

11 Para que tudo isso serve?

Eis a seção mais difícil de escrever desse texto. Não por falta de exemplos, mas pelo excesso deles. O que apresentei nessas poucas linhas são a caixa de ferramentas fundamentais de qualquer matemático ou físico. Garanto que, todo dia, um matemático usa mais coisas desse texto que somas ou multiplicações. Sou físico, não matemático, então vou mandar aqui alguns exemplos de física.

Na mecânica quântica, os estados físicos não são mais descritos por pontos, massas ou bloquinhos, você precisa tratar tudo com mais calma e mais matemática. Por isso, logo de cara, dizemos que os sistemas físicos são elementos de um espaço de Hilbert. Esse espaço você agora é capaz de entender o que é, um espaço de Hilbert é um espaço métrico munido de um produto interno nos complexos que possui a propriedade de ser completo (algo que não citei, mas que você facilmente consegue achar, ler e entender). Toda vez que lidamos com mecânica quântica, precisamos ter em mente onde moram os elementos que calculamos, quais as propriedades desse espaço e quais teoremas podem nos ajudar.

Dou um exemplo concreto. Recentemente eu estudava o livro "Operator Algebras and Quantum Statistical Mechanics", de Bratteli e Robinson, que é um livro razoavelmente avançado de física estatística. Logo no começo, enquanto ele dissertava sobre os fundamentos do que ia falar, coloca a frase:

> Assumamos que os estados de cada partícula formem um espaço de Hilbert complexo $\mathfrak{h}$ e tomemos $\mathfrak{h}^n = \mathfrak{h} \otimes \mathfrak{h} \otimes \ldots \otimes \mathfrak{h}$. Efetuando a soma direta entre todos os possíveis espaços $\mathfrak{h}^n$, teremos $\bigoplus_n \mathfrak{h}^n = \mathfrak{F}(\mathfrak{h})$. Todo vetor de $\mathfrak{F}(\mathfrak{h})$ é um sequência de vetores de $\mathfrak{h}^n$, e $\mathfrak{h}^n$ pode ser identificado a um subespaço fechado de $\mathfrak{F}(\mathfrak{h})$ formado pelo vetores com todos os componentes exceto o n-ésimo nulos.

Note que, apesar de não ser muito compreensível fora do contexto da quântica, nenhum dos termos desse parágrafo é estranho a esse texto. Ele fala de espaço de Hilbert complexo, sequência de vetores e espaços fechados. Com isso, você está munido de um produto interno em um espaço vetorial complexo completo e pode aplicar qualquer um dos teoremas desse texto nele, se a ocasião chamar.

Na relatividade geral, nossa noção de espaço tempo já era. O que mudamos na relatividade geral não são exatamente as leis da física, mas a distância entre dois elementos. Antes, dizíamos que a norma entre dois pontos no espaço era

$$d_2(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2},$$

mas isso não será mais verdade. Vamos mudar a norma do espaço e fazer com que a norma dependa das coordenadas: a maneira como você mede distâncias dependerá do que está dentro

do espaço medido (essa é aquela história de “distorção do espaço tempo”). O produto interno também mudará, o ângulo entre dois vetores também, mas você sabe calcular essas coisas e tem teoremas que funcionam para qualquer espaço, qualquer produto interno. Não importa o quão estranha ou deformada seja a modificação no espaço tempo, as normas ou os produtos internos, você sempre terá a desigualdade de Cauchy-Schwarz.