

Números Duais

Alexandre Luís Kundrát Eisenmann

Maio 2019

Uma ideia ousada

Em algum momento do século XVI, um italiano foi iluminado por uma ideia ingênua, quase infantil, mas ousada. E por que não definir um número que, multiplicado por ele próprio, é igual a -1 ? Como todos sabemos, números multiplicados por eles próprios resultam em valores positivos, afinal "menos com menos dá mais", não é mesmo?

Girolamo Cardano foi o primeiro a sugerir que raízes de números negativos possam, talvez, ter um papel na matemática. Ao analisar a equação abaixo, notou que se ignorássemos as "torturas mentais envolvidas", a raiz obtida pela fórmula de Báskara é de fato a solução para o problema.

$$x(10 - x) = 40 \Rightarrow x^2 - 10x + 40 = 0 \Rightarrow x = 5 \pm \sqrt{-15}$$

Substituindo x por $5 + \sqrt{-15}$, Cardano observou que

$$x(10 - x) = (5 + \sqrt{-15})(5 - \sqrt{-15}) = 5^2 - (\sqrt{-15})^2 = 25 - (-15) = 40$$

O próximo e mais significativo passo dado foi por outro italiano, Rafael Bombelli, desta vez trabalhando com a resolução das equações cúbicas. Postulando a existência do número $\sqrt{-1}$, Bombelli resolveu uma equação cúbica cuja resolução esbarrava nestes números impossíveis. O interessante foi que em algum momento do desenvolvimento, os números estranhos se cancelaram e a resposta obtida foi um número real, facilmente verificável, indicando que o método seguido estava no caminho certo.

Não detalharei aqui o feito de Bombelli, a história das equações cúbicas é bem conhecida e digna de um roteiro de Hollywood, recomendo a leitura.

Uma ideia heterodoxa e ousada mudou a história da matemática para sempre, os números complexos, como os conhecemos, ganharam status de números de

fato e possuem hoje em dia inúmeras aplicações. Na perspectiva que temos hoje, é possível ver a simplicidade e beleza desta pequena ideia:

$$i^2 = -1$$

Meio milênio depois conto a história de Ben Christopher. Ben é australiano e no momento dos acontecimentos que serão descritos, trabalhava, assim com eu, em uma grande empresa pública da Austrália como analista de negócios. Ben estava há alguns meses de sua graduação em matemática, mas ele já era graduado em física e numa rápida conversa era evidente que Ben possuía uma inteligência aguda e fora do comum.

Sabendo de minha formação em matemática, Ben encontrou alguém para conversar. Sim, nós matemáticos somos carentes e achar alguém para conversar sobre matemática é meio caminho andado para uma nova amizade. No começo as conversas não eram fáceis, Ben falava rápido e com profundidade, precisei voltar a estudar alguns temas para poder acompanhá-lo. Em algum momento, brincando, comecei a pedir a bibliografia para me preparar para uma nova sessão de conversa informal no cafezinho da empresa.

Certo dia ele resolveu confidenciar a pesquisa que estava fazendo. Uma ideia ingênua, quase banal, e extremamente poderosa. Ben apresentou-me um novo número, batizado por ele de *sero number*. Este estranho número era definido como um número cujo quadrado era igual a zero.

$$s^2 = 0$$

Imediatamente percebi a similaridade com os números complexos. Um número que multiplicado por si mesmo resulta em zero. Claro, inicialmente fiquei confuso, dizendo que este número deveria ser igual a zero. Ben, antecipando meu comentário, logo detalhou sua definição informando que o número *sero* é, por definição, diferente de zero porém seu quadrado é zero.

A intuição por trás deste número, ele explicava, é de que ele era tão pequeno, mais tão pequeno que quando elevado ao quadrado ele se tornava zero, daí o nome, a palavra zero adulterada como um "s" na frente. Vale notar que quando um número é menor do que 1, seu quadrado é menor do que o número original. Exemplo: $(\frac{1}{10})^2 = \frac{1}{100}$

Calculando derivadas

Ben foi ousado, mostrou que através dos números *sero*, podemos calcular a derivada de uma curva sem usar o complexo conceito de limites, sendo necessário

apenas a intuição da reta tangente.

Sabemos, por exemplo, que a derivada da função $f(x) = x^2$ é igual a $2x$. Classicamente esta, e todas as outras regras de derivação, são obtidas resolvendo a seguinte expressão:

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

Onde f é a função que estamos estudando. A intuição é de que quanto menor é a variável h , mais a mais o quociente acima se aproxima da inclinação da reta tangente no ponto x .

Quando os estudantes são apresentados ao conceito de limite pela primeira vez, nem tudo parece tão simples. A notação não reflete de forma tão direta as operações que deverão realizar e os detalhes e nuances afugentam boa parte dos alunos.

Felizmente, a ideia de Ben vem a nosso resgate. Vamos utilizar o número *sero* para representar um número minúsculo e esquecer a ideia de limites por hora. A expressão acima, convertida, se parece com:

$$\frac{f(x+s) - f(x)}{s} \tag{1}$$

Onde f é a função que estamos estudando e s número *sero*. Vamos agora encontrar a derivada para a função $f(x) = x^2$ resolvendo a expressão acima:

$$\frac{(x+s)^2 - x^2}{s} = \frac{x^2 + 2xs + s^2 - x^2}{s} = \frac{2xs + \overset{0}{\cancel{s^2}}}{s} = 2x$$

Voilà, quando usamos o fato de que s ao quadrado é zero, eliminamos esta parcela e o resultado é simplificado para $2x$. Ben mostrou que todas as derivações elementares podem ser obtidas por essa simples ideia sem recorrer aos limites, incluindo a logaritmos que mostrarei abaixo e as funções trigonométricas.

Não é difícil reorganizar a expressão acima para o caso mais geral onde $f(x) = x^n$ e deixo para o leitor obter sua derivada.

Ben também mostrou algumas identidades interessantes, entre elas o valor do número e^x . Sabemos que a expansão de Taylor para o número e^x é dada por

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Substituindo x por s

$$e^s = 1 + s + \cancel{\frac{x^2}{2!}}^0 + \cancel{\frac{x^3}{3!}}^0 + \dots = 1 + s$$

E, analogamente, substituindo x por xs

$$e^{xs} = 1 + xs \quad (2)$$

Desenvolvendo a expressão (1) substituindo f pela função \ln é possível calcular a derivada da função $f(x) = \ln(x)$ que sabemos ser igual a $\frac{1}{x}$.

$$\begin{aligned} \frac{\ln(x+s) - \ln x}{s} &= \frac{\ln(x(1 + \frac{s}{x})) - \ln x}{s} \\ &= \frac{\cancel{\ln x} + \ln(1 + \frac{s}{x}) - \cancel{\ln x}}{s} \\ &= \ln \frac{1 + \frac{s}{x}}{s} \quad (\text{usando equação (2)}) \\ &= \frac{\ln e^{\frac{s}{x}}}{s} \\ &= \frac{s/x}{s} = \frac{1}{x} \end{aligned}$$

Após a análise destes e de outros casos particulares, Ben partiu para discussões mais gerais envolvendo a série de Taylor. Assumindo que uma função f pode ser expandida por uma série de Taylor no ponto a conforme a expressão

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \dots$$

Ben notou que o valor de $f(a+s)$, onde s é o número *sero* é igual a

$$f(a+s) = f(a) + \frac{f'(a)}{1!}(s) + \cancel{\frac{f^{(2)}(a)}{2!}(s)^2}^0 + \dots = f(a) + f'(a)s \quad (3)$$

Magicamente a equação (3), quando isolamos $f'(a)$, é exatamente igual a definição de derivada que apresentamos com expressão 1. Neste ponto fiquei perplexo, quando *plugamos* um número $x+s$ numa função diferenciável obtemos o valor

$f(x)$ somado a sua própria derivada multiplicada pelo número *sero*. Além disso, de quebra, derivamos (perdoem o trocadilho) a própria definição da derivada.

Lembro bem o excitamento que senti nestas semanas enquanto aprendia, discutia e revia o trabalho de Ben. Era um privilégio poder assistir o desenrolar do nascimento de uma ideia matemática original e poderosa.

Números Duais

Em meados de 2016, época onde os fatos aqui narrados se desenvolveram, eu estava analisando alguns modelos de *machine learning* na empresa em que Ben e eu trabalhávamos. Curioso com o funcionamento do algoritmo de otimização *gradient descent* achei um pequeno artigo que detalhava como a relativamente nova linguagem de programação *Julia* efetuava a chamada diferenciação automática.

Nas entranhas do artigo havia a referência a uma estrutura algébrica chamada *Dual Numbers* ou Números Duais em português. Parecia que esses estranhos números eram de fato importantes para o desenvolvimento do artigo e em algum ponto resolvi entender mais a fundo do que se tratavam.

Logo na primeiro parágrafo da wikipedia sobre *Dual Numbers* em inglês, eu notei uma expressão familiar, desta vez com a letra grega épsilon $\epsilon^2 = 0$. À medida que me aprofundava na leitura mais eu tinha certeza do que se tratava. Em algum ponto, fui convencido, os tais dos Números Duais eram de fato os *Sero numbers*.

A formulação dos números Duais na wikipedia seguia uma exposição similar ao comumente usado para os números complexo. Um número Dual foi definido como um número da forma $a + b\epsilon$ onde ϵ possui a propriedade $\epsilon^2 = 0$. O número *sero* na formulação de Ben era, portanto, equivalente ao componente ϵ do número Dual.

À exemplo dos números complexos, se p e q são números duais $p = x + y\epsilon$ e $q = u + v\epsilon$, podemos definir as 4 operações básicas abaixo. Note que a divisão só pode ser definida quando $u \neq 0$.

$$\begin{aligned}
p + q &= (x + y\epsilon) + (u + v\epsilon) = (x + u) + (y + v)\epsilon \\
p - q &= (x + y\epsilon) - (u + v\epsilon) = (x - u) + (y - v)\epsilon \\
p \times q &= (x + y\epsilon)(u + v\epsilon) = xu + (xv + yu)\epsilon \\
p \div q &= \frac{x + y\epsilon}{u + v\epsilon} = \frac{xu + (uy - xv)\epsilon}{u^2}
\end{aligned}$$

O artigo também expunha um dos principais resultados obtidos por Ben, a fórmula (3) que repito abaixo usando a notação dos números duais.

$$f(a + \epsilon) = f(a) + f'(a)\epsilon$$

Este resultado é importante e vou insistir um pouco em uma simples aplicação. Suponha que desejássemos encontrar a derivada da função $f(x) = 2x^2 + x + 5$ no ponto $x = 3$. Aplicando a fórmula anterior com um número Dual cuidadosamente selecionado, $3 + \epsilon$, obteremos não só o valor da derivada da função no ponto $x = 3$ como também o valor da função no mesmo ponto.

$$f(3 + \epsilon) = 2(3 + \epsilon)^2 + (3 + \epsilon) + 5 = 2(9 + 6\epsilon + \overset{0}{\cancel{\epsilon^2}}) + 8 + \epsilon = 26 + 13\epsilon$$

Notem que no número dual $26 + 13\epsilon$ encontramos não só valor da função no ponto 3, $f(3) = 26$ com também, simultaneamente, sua derivada no mesmo ponto, $f'(3) = 13$

Eu não estava feliz quando escrevi um email para Ben com um breve *Have a look at this* com o específico link da wikipedia. Antecipando seu desapontamento, considerei não comentar o assunto, e por alguns minutos flertei com a ideia de não apertar o botão *Send*. Finalmente fui em frente, racionalizando que era a coisa certa a fazer.

Machine Learning

Não é a primeira vez na matemática que um resultado teórico obtido na matemática pura encontre aplicações centenas de anos mais tarde. Os números complexos

são um exemplo clássico, criado na Itália renascentista encontrou diversas aplicações, em especial na engenharia elétrica.

Já os números Duais encontraram aplicação na diferenciação automática e, por consequência, são peças fundamentais no arcabouço tecnológico que possibilita o aprendizado de máquinas ou *Machine Learning* em inglês, portanto protagonistas na assustadora revolução que ocorre diante de nossos olhos.

Um computador "aprende" como solucionar um problema "escolhendo" quais são os parâmetros de uma determinada função (*weights*) que minimizem o erro do resultado obtido com os dados reais (*training set*). A "escolha" deste parâmetros é um clássico problema de otimização que pode ser resolvido da forma "deriva-se a função, iguala-se a zero e resolve a equação".

Quando o número de parâmetros é enorme, não é possível uma solução analítica, mas sim um processo iterativo onde o ponto de mínimo é alcançado passo a passo, num algoritmo chamado *gradient descend*, onde a direção a ser seguida é apontada pelo gradiente da função neste espaço multidimensional.

As facilidades operacionais possibilitadas pelos números Duais, tornam esses números ideais para a implementação da *forward propagation* numa linguagem de programação, etapa em que os gradientes (derivada em múltiplas dimensões) são calculados.

Ben Christopher

Ben Christopher é uma pessoa real, assim como os acontecimentos descritos neste ensaio. Tenho certeza de que seu trabalho ocorreu independentemente do trabalho dos números Duais, fenômeno não raro na história da matemática. Um dos casos mais famosos foi o próprio cálculo diferencial, desenvolvido independentemente por Isaac Newton e Gottfried Wilhelm Leibniz.

Como observador próximo, me senti privilegiado por testemunhar de perto a expressão da criatividade e força de uma ideia matemática. Em pleno século XXI, a matemática está mais viva do que nunca, o que será que os anos vindouros irão nos preparar?

No dia seguinte após o recebimento meu desagradável email, Ben estava visivelmente abatido e cabisbaixo. Ele verbalizou seu despondimento, afinal seu trabalho acabara de perder o ineditismo precedido pelo trabalho do matemático William Clifford em 1873.

Passado algumas semanas, recuperou seu bom humor e energia, devorando tra-

balhos sobre os números Duais e iniciando investigações em novas áreas. Ben havia estendido a aplicação de seus números para a integração definida e indefinida, mares, aparentemente, não antes navegados.