Derivação automática

Recap e aproximação por diferenças finitas

Nemo

July 2, 2024

1 Objetivos

Essa serie vai tentar explicar os conceitos e motivações por trás da ideia conhecida por "derivação automática". Começaremos por uma revisão das ideias básicas de calculo (de modo breve), veremos como "computar" aproximações numéricas e expressões simbólicas da derivada de uma função, e as limitações e problemas desses métodos. Finalmente então implementaremos uma pequena biblioteca do que é conhecido por diferenciação automática "Forward mode".

1.1 Nosso plano

1. Introdução
2. Recap de calculo
3. Implementar diferenças finitas
4. (talvez) implementar diferenciação simbólica
5. Base teorica do autodiff
6. Implementação de autodiff

2 Recapitulando: Derivação

O quanto uma função $F(x) = y$, com $x, y \subset \mathbb{R}$, muda conforme x varia é chamado de derivada.

A velocidade é um exemplo claro de como a derivação representa a mudança. Se a posição de um carro que se move é descrita por $p(t)$, sua velocidade vai ser a derivada da posição em função de t e sua aceleração a derivada da derivada (a variação da variação), ou em "termos matemáticos":

$$Velocidade = v(t) = \frac{dp(t)}{dt} \tag{1}$$

$$Acelerao = a(t) = \frac{dv(t)}{dt} = \frac{d^2p(t)}{d^2t} \tag{2}$$

A derivada de função $F(x)$ é formalmente definida utilizando limites:

$$\frac{dF(x)}{dt} = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} \tag{3}$$

Para uma função de uma dimensão, graficamente isso pode ser visto como tentar encontrar a reta tangente a um ponto especifico, ou em outras palavras: uma aproximação linear daquela função.

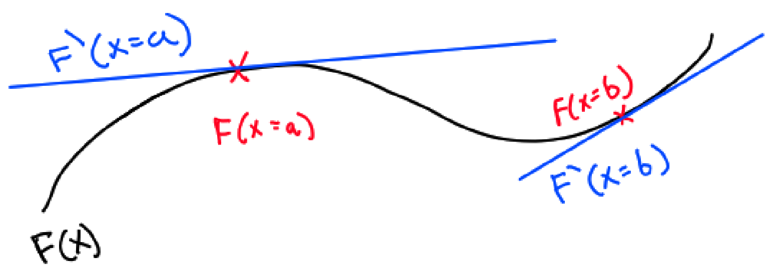


Figure 1: Derivada de $F(x)$ com $x = a$ e $x = b$

Utilizando as propriedades de limites é possível determinar o comportamento para maioria das operações realizadas nos espaços "normais", é dai que vem as tabelas de derivação. As tabelas de derivação apresentam um conjunto de **regras simbólicas** para manipular essas expressões, como por exemplo as presentes na Tabela 1.

Table 1: Algumas regras de derivação simbólica

Comentário	$Expr$	$\frac{d(Expr)}{dx}$	Exemplo
Regra da adição	$F \pm W$	$F' \pm W'$	$x^3 - x^2 \Rightarrow 3 \cdot x^2 - 2 \cdot x$
Regra do produto	$F * W$	$F' \cdot W + W' \cdot F$	$e^x \cdot x^2 \Rightarrow e^x \cdot x^2 + 2 \cdot x \cdot e^x$
Regra da cadeia	$F(W(x))$	$F'(G(x)) \cdot G'(x)$	$sin(x) \circ x^2 \Rightarrow \ cos(x^2) \cdot 2 \cdot x$
Regra para reciproca	$\frac{1}{F}$	$\frac{-F'}{F^2}$	$\frac{1}{x} \Rightarrow -\frac{1}{x^2}$
Regra para polinômios	x^n	$n \cdot x^{n-1}$	$x^5 \Rightarrow 5 \cdot x^4$
Constantes, números	K	0	$42 \Rightarrow 0$
...	...	...	...

2.1 Notação

Calculo possui uma historia um pouco complicada com notações. Existem varias, e dependendo da situação convém usar uma ou outra. Todas são um pouco confusas. A derivada de uma função $F(x)$ pode ser representada:

- Com um apóstrofo ou similar: $F'(x)$
- Com um circulo sobre o nome da função: $\dot{F}$
- Como um operador D(iferencial): $DF(x)$
- Como uma especie de razão entre a função e um "infinitesimal": $\frac{dF}{dx}$

Quase todas podem ser estendidas para tratar de derivadas de ordem maior e/ou derivadas parciais, que não cobriremos aqui.

Para saber mais:

- Introdutório : `https://www.youtube.com/watch?v=6v0SMTZ8hkU`
- Introdutório : Um Curso de Cálculo - Vol. 1, Hamilton Luiz Guidorizzi
- Avançado / rigoroso : Calculus on Manifolds, Michael Spivak, 1965

2.2 Diferenças finitas

Para de fato calcularmos essa função em um ponto utilizando computadores, podemos utilizar uma aproximação. Isso é feito dando valores pequenos para h em vez de usar um limite. Essa aproximação então é:

$$F'(x) \approx \frac{F(x+h) - F(x)}{h} \tag{4}$$

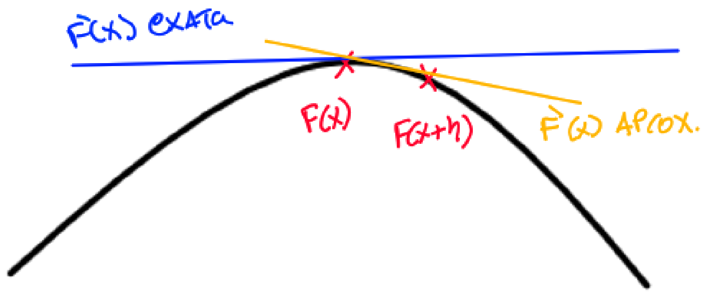


Figure 2: Aproximação por diferenças finitas "para frente" de $F'(x)$.

A equação anterior aproxima a derivada como um "passo" para frente, mas é possível também aproximá-la como um passo para trás:

$$F'(x) \approx \frac{F(x) - F(x-h)}{h} \tag{5}$$

Ou meio passo para frente e meio para trás:

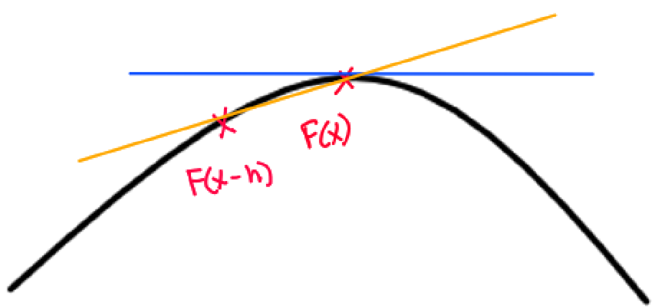


Figure 3: Aproximação por diferenças finitas "para trás" de $F'(x)$.

$$F'(x) \approx \frac{F(x+\frac{h}{2}) - F(x-\frac{h}{2})}{h} \tag{6}$$

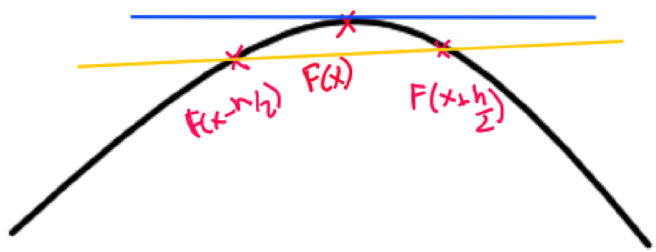


Figure 4: Aproximação por diferenças finitas "centrada" de $F'(x)$.

As formulas que envolvem somente uma direção se comportam de modo similar em precisão e estabilidade, mas a diferença "centrada" tem um erro menor.

2.3 Implementação de diferenças finitas

Primeiramente convém definir algumas funções que vão nos auxiliar a extrair e plotar os resultados.

2.3.1 Funções auxiliares

```
(define (range from by to)
  ;; Função que gera um intervalo numerico de [from] em passos [by] até [to].
  (if (>= from to)
      '()
      (cons from (range (+ by from) by to))))

(define (log10-range start end n-per-decade)
  ;; função que gera um intervalo de [n-per-decade] pontos espaçados de modo uniforme
  ;; em uma decada, em (- [decades] [start]) decadas.
  (let ([decades (range start 1 end)]
        [points (range 1 (/ 10 n-per-decade) 10)])
    (apply append
           (map (lambda (decade)
                  (map (lambda (point)  (* point (expt 10.0 decade))) points))
```

```
                decades))))

(define (relative-error a b)
  ;; calcula o erro relativo entre duas coisas. err = (abs (a - b) / b)
  (abs (/ (- a b) b)))


(define (print-row-with-spaces lst)
  ;; imprime uma lista [lst] com seus elementos separados por " " e termina com \n.
  (format #t "~{~a~^ ~}~%" lst))

(define (print-dataset . list-of-lists)
  ;; imprime o dataset todo
  (apply  for-each (lambda x (print-row-with-spaces x))  list-of-lists))
```

2.3.2 Função de exemplo

Para demonstrar as formas de se calcular a aproximação por diferenças finitas, convém escolher uma função. Escolhi, por nenhuma razão especifica, utilizar um seno amortecido:

$$\sin(x) \cdot e^{\frac{-x}{5}} \tag{7}$$

Que tem como derivada:

$$-\frac{\mathrm{e}^{-\frac{x}{5}}\left(\sin(x) - 5\cos(x)\right)}{5} \tag{8}$$

Suas definições em scheme são representadas abaixo.

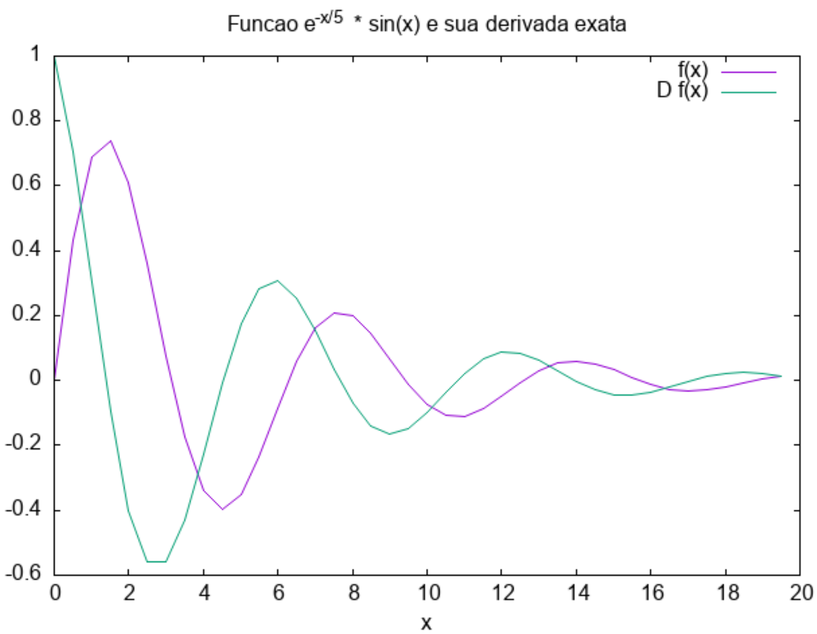
```
(define (my-test-func x)
  ;; e^(-x/5) * sin(x)
  (* (sin x) (exp (/ (* -1 x) 5))))

(define (my-test-derivative x)
  ;;-(e^(-x / 5) * (sin(x) - 5 * cos(x))) / 5
  (* -1 (/ (* (exp (/ (* -1 x) 5)) (- (sin x) (* 5 (cos x)))) 5)))
```

Os dois blocos de codigo abaixo gravam e leem os resultados dessa função e sua derivada em um arquivo, e plotam o resultado para $[0, 20]$. O processo de escrita no arquivo **é realizada pelo org-mode** que executa o codigo scheme, coleta a saida, e salva. Já a leitura e plotagem são feitas no gnuplot (voce também pode usar python, R ou sua linguagem favorita para plotar).

```
(let* ([dominio (range 0 0.5 20)]
       [plot-test-func  (map my-test-func dominio)]
       [plot-test-Dfunc (map my-test-derivative dominio)])
  (print-dataset dominio plot-test-func plot-test-Dfunc))
```

```
# Script gnuplot
reset
set title "Funcao {e^{-x/5}}  * {sin(x)} e sua derivada exata"
set xlabel "x"
#plota a as duas funções
plot data u 1:2 w l title "f(x)", data u 1:3 w l  title "D f(x)"
```



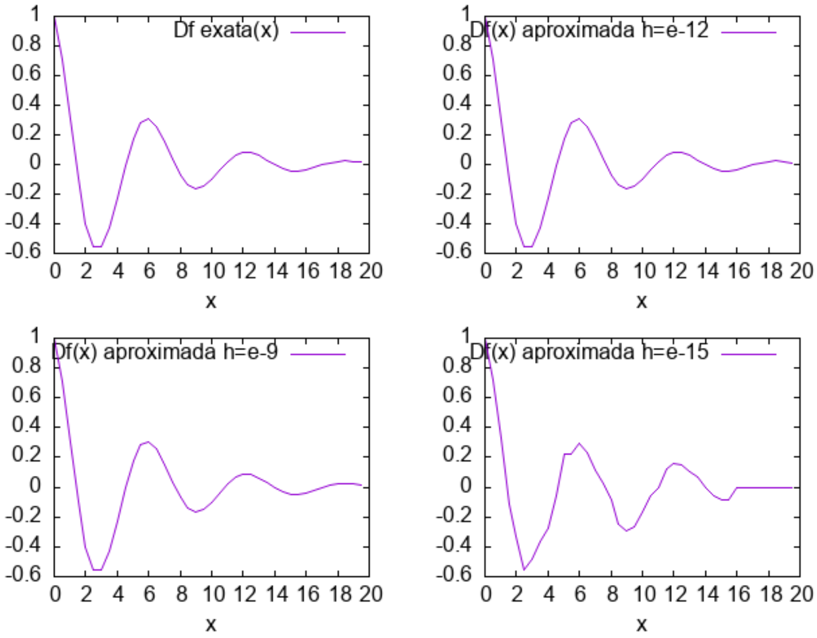
O bloco abaixo plota a derivada aproximada, utilizando a aproximação "para frente", para três valores de h: 1e-9,1e-12 e 1e-15. É possivel ver que a aproximação para de funcionar para o menor valor, e se continuassemos a diminuir h ela quebraria completamente.

```
(define (aprox-d f h x)
    (/ (- (f (+ x h)) (f x)) h))
```

```
(let* ([dominio (range 0 0.5 20)]
       [plot-exact-Dfunc (map my-test-derivative dominio)]
       [plot-aprox-Dfunc-e9  (map (lambda (x) (aprox-d my-test-func 1e-9 x)) dominio)
       [plot-aprox-Dfunc-e12 (map (lambda (x) (aprox-d my-test-func 1e-12 x)) dominio
       [plot-aprox-Dfunc-e15 (map (lambda (x) (aprox-d my-test-func 1e-15 x)) dominio
  (print-dataset dominio plot-exact-Dfunc plot-aprox-Dfunc-e9 plot-aprox-Dfunc-e12 pl

# Script gnuplot
reset

set xlabel "x"
set multiplot layout 2,2 columns
#plota a as duas funções
plot data u 1:2 w l title "Df exata(x)"
plot data u 1:3 w l  title "Df(x) aproximada h=e-9"
plot data u 1:4 w l  title "Df(x) aproximada h=e-12"
plot data u 1:5 w l  title "Df(x) aproximada h=e-15"
unset multiplot
```



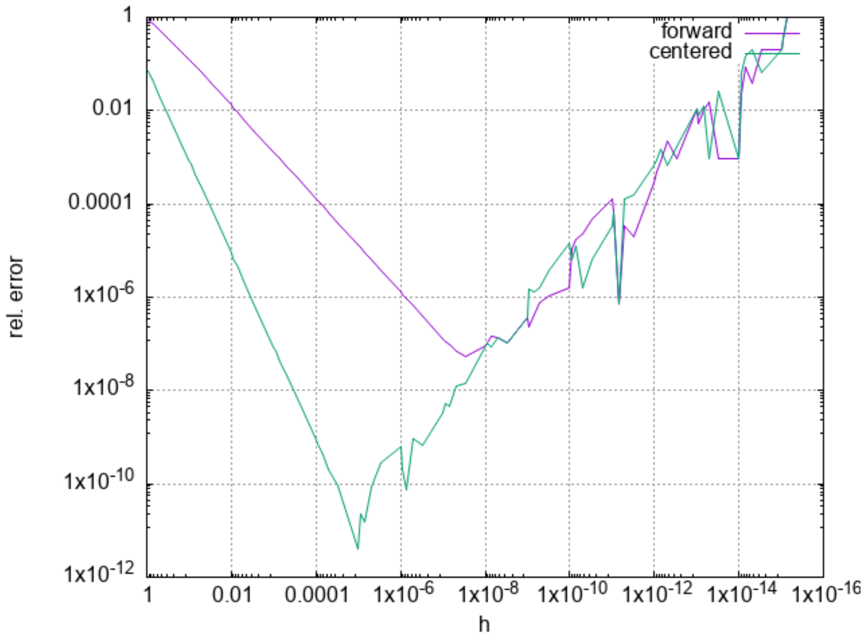
Para observar o comportamento do erro, em relação a derivada "ideal", convém selecionarmos um ponto, neste caso $x = 8$, e variarmos h e o tipo de aproximação. Os codigos abaixo fazem isso para os três jeitos de aproximar que vimos anteriormente.

É possível observar que o erro relativo diminui de maneira linear para todos os três, até um certo ponto, depois disso ele começa a subir de um jeito errático. Também é possível ver que o erro da aproximação centrada é consideravelmente menor. O erro após a "quina" decorre de alguns fatores, mas principalmente de arredondamento e de "formula" (pois é uma aproximação).

```
(define (aprox-d-centered f h x)
  (/ (- (f (+ x (/ h 2))) (f (- x (/ h 2)))) h))

(let* ([dominio-h (log10-range -16 0 5)]
       [x 8]
       [derivada-em-x (my-test-derivative x)]
       [aproximacao-em-x (map (lambda (h) (aprox-d my-test-func h x)) dominio-h)]
       [aproximacao-em-x-centrada (map (lambda (h) (aprox-d-centered my-test-func h x)
       [rel-error-f (map (lambda (i) (relative-error i derivada-em-x)) aproximacao-em-
       [rel-error-c (map (lambda (i) (relative-error i derivada-em-x)) aproximacao-em-
  (print-dataset dominio-h rel-error-f rel-error-c ))

# Script gnuplot
reset
set grid
set xlabel "h"
set ylabel "rel. error"
set logscale x 10
set logscale y 10
set xrange reverse
#plota a as duas funções
plot data u 1:2 with line title "forward",\
     data u 1:3 with line title "centered"
```



forward
centered
rel. error
h
1
0.01
0.0001
1x10-6
1x10-8
1x10-10
1x10-12
1x10-14
1x10-16