

Trabajo integrador: Diferenciación automática vs. diferencias finitas

Dr. Ing. Benjamin A. TOURN - Ing. Carlos G. MASSOBRIO

Martes 22 de Octubre - B52024

1 Enunciado

Diferenciación automática

La diferenciación automática (*autodiff*) [1] es una metodología que permite calcular de manera elegante las derivadas parciales de cualquier función arbitraria en un punto dado. Esencialmente, el mecanismo *autodiff* descompone la función en una secuencia de operaciones aritméticas elementales (+, −, *, /) y funciones (máx, exp, log, cos, sin, etc.); luego utiliza la regla de la cadena para computar la derivada de la función original con respecto a sus parámetros iniciales.

Reverse-mode autodiff

Considere la función

$$f(\mathbf{x}) = 8(x_1^3 + x_2 x_3), \quad (1.1)$$

donde $\mathbf{x} = (x_1, x_2, x_3)$. El grafo computacional correspondiente a esta función viene dado en la figura 1:

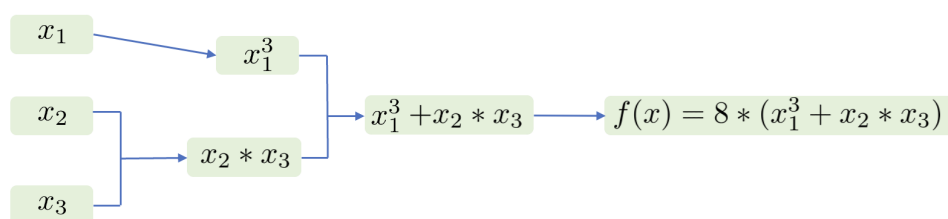


Figura 1: Grafo computacional para la función $f(\mathbf{x})$.

Reverse-mode incluye dos pasos principales: *forward* y *backward*. Durante el paso *forward* las entradas son pasadas aguas abajo a través del grafo tal como se muestra en el esquema 2. Para ayudar con la comprensión, se ha asignado una nueva variable para cada operación tal como puede observarse en las expresiones resaltadas en azul.

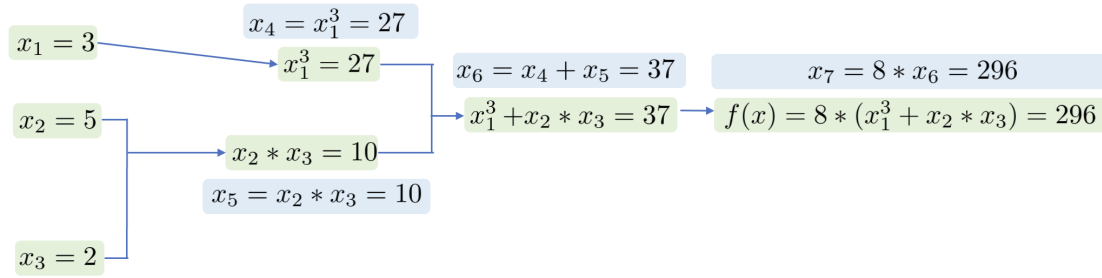


Figura 2: Paso *forward* para el grafo computacional de la función $f(x)$.

En el paso *backward*, se utiliza la regla de la cadena para calcular los valores del gradiente de la función $f(x)$ con respecto a x_1, x_2, x_3 , como se muestra en el esquema 3. Luego, recorriendo el grafo de derecha a izquierda, podemos calcular las derivadas parciales de f respecto a las entradas. Dichas derivadas se computan de manera local, es decir que al modificar los valores en las entradas se modificará el valor de las derivadas.

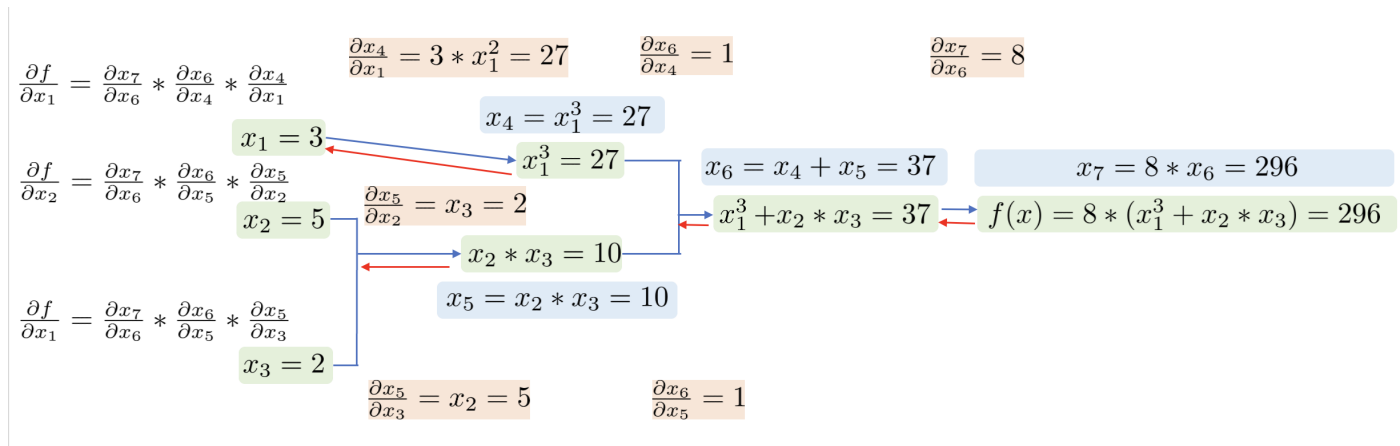


Figura 3: Paso *backward* para el grafo computacional de la función $f(x)$.

Método de diferencias finitas

Considere la definición de la derivada de una función f en el punto x :

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}. \quad (1.2)$$

El cociente incremental $(f(x+h) - f(x))/h$ es una aproximación de la derivada $f'(x)$, que se torna más precisa a medida que h disminuye. Para conocer que tan lejana es esta estimación respecto al valor verdadero, se utiliza la expansión en serie de Taylor:

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(\eta), \quad \eta \in [x, x+h]. \quad (1.3)$$

Al reescribir la expresión anterior como:

$$\varepsilon = \frac{f(x+h) - f(x)}{h} - f'(x) = \frac{h}{2}f''(\eta), \quad (1.4)$$

es posible ver que el error ε es proporcional a h^1 . Por lo tanto, $(f(x+h) - f(x))/h$ brinda una aproximación de primer orden de $f'(x)$. Si $h = \Delta x$ es un número finito positivo, entonces $(f(x + \Delta x) - f(x))/\Delta x$ es conocido como aproximación de primer orden $O(\Delta x)$ hacia adelante de $f'(x)$. Por su parte, $(f(x) - f(x - \Delta x))/\Delta x$ constituye la aproximación de primer orden hacia atrás de $f'(x)$ obtenida con $\Delta x = -h$. A su vez, $f'(x)$ puede ser aproximada mediante el cociente incremental simétrico $(f(x + \Delta x/2) - f(x - \Delta x/2))/\Delta x$. La expansión de serie de Taylor para este caso, luego de operar algebraicamente (¡comprobar!), conduce a la siguiente expresión:

$$\frac{f(x + \Delta x/2) - f(x - \Delta x/2)}{\Delta x} = f'(x) + \frac{\Delta x^2}{24} f'''(\eta), \quad \eta \in [x - \Delta x/2, x + \Delta x/2], \quad (1.5)$$

que constituye una aproximación de segundo orden $O(\Delta x^2)$ para $f'(x)$. De esta manera, es posible utilizar diferentes esquemas para aproximar la misma derivada, con diferentes órdenes de precisión.

En conclusión, la aproximación de las derivadas de una función mediante las estrategias descritas constituye la esencia del Método de Diferencias Finitas (FDM, por sus siglas en inglés), y puede ser empleado para resolver ecuaciones en derivadas parciales [2]. A continuación se muestran dos de las expresiones mayormente empleadas de las aproximaciones de segundo y cuarto orden de las derivadas primera y segunda, respectivamente:

$$f'(x) : \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x} \quad (1.6a)$$

$$f''(x) : \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x))}{\Delta x^2}. \quad (1.6b)$$

2 Actividades

1. Resolver las ecuaciones [3]:

$$\nabla^2 u = \sin(\pi x) \sin(\pi y); \quad 0 < x < 1, 0 < y < 1; \quad (2.1a)$$

$$u(0, y) = u(1, y) = 0; \quad (2.1b)$$

$$u(x, 0) = u(x, 1) = 0, \quad (2.1c)$$

mediante PINN y mediante el método de diferencias finitas en grillas de tamaños 5×5 , 10×10 y 20×20 . Para cada tamaño de grilla utilizar una configuración de red neuronal de $[2, 3, 3, 1]$, $[2, 5, 5, 1]$ y $[2, 10, 10, 1]$. Comparar ambas soluciones con la solución exacta dada por:

$$u(x, y) = -1/(2\pi^2) \sin(\pi x) \sin(\pi y) \quad (2.2)$$

¿Qué solución es más precisa? ¿Por qué? Repita el experimento utilizando otras arquitecturas.

2. Resolver el problema de conducción de calor con término fuente no lineal dado por [3]:

$$\nabla^2 u = 0.5e^u; \quad 0 < x < 1, 0 < y < 1; \quad (2.3a)$$

$$u(0, y) = u(x, 0) = 0; \quad (2.3b)$$

$$\frac{\partial u(1, y)}{\partial x} = \frac{\partial u(x, 1)}{\partial y} = 0, \quad (2.3c)$$

mediante PINN y mediante el método de diferencias finitas en grillas de tamaños 5×5 , 10×10 y 20×20 . La configuración de la red neuronal debe ser $[2, 5, 5, 1]$. Analice y discuta los resultados. Repita el proceso, pero utilizando muestreos aleatorios de los puntos de colocación con la misma cantidad de muestras que en el caso de las grillas uniformes. ¿Qué observa?

Referencias

- [1] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [2] Aniruddha Bora and Weizhong Dai. Gradient preserved method for solving heat conduction equation with variable coefficients in double layers. *Applied Mathematics and Computation*, 386:125516, 2020.
- [3] MWM Gamini Dissanayake and Nhan Phan-Thien. Neural-network-based approximations for solving partial differential equations. *communications in Numerical Methods in Engineering*, 10(3):195–201, 1994.