Vector gradiente, derivada direccional y matriz Hessiana

Gamaliel Moreno
Maestría en Ciencias del Procesamiento de la Información

UAZ

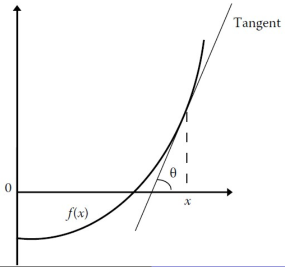
Agosto-Diciembre, 2021

Vector gradiente

La derivada o el gradiente de una función $f(x)$ en un punto x, generalmente denotada como $f'(x)$, es la pendiente de la tangente en el punto.

$$f'(x) = tan(\theta)$$

donde θ es el angulo medido por la tangente respecto a la horizontal.



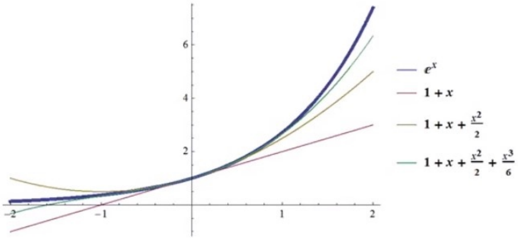
Derivada y gradiente

A lo largo de la dirección gradiente, hay un máximo cambio en el valor de la función. Esto es , el gradiente da información necesaria para localizar el máximo o mínimo de una función. En la mayoría de los problemas de optimización la $f'(x)$ tiene que ser evaluada numéricamente.

- Derivada adelantada
- Derivada atrasada
- Derivada central

Aproximaciones de la serie de Taylor

Una función puede aproximarse por medio de una serie de polinomios.



Su aproximación consiste en elegir adecuadamente los coeficientes.

Aproximaciones de la serie de Taylor

La serie de Taylor de una función f real o compleja $f(x)$ infinitamente diferenciable en el entorno de un número real o complejo a esta dada por

$$f(x) = f(a) + \frac{f^{'}(a)}{1!}(x-a) + \frac{f^{''}(a)}{2!}(x-a)^2 + \frac{f^{'''}(a)}{3!}(x-a)^3 + \cdots$$

$$+\frac{f^{N}(a)}{N!}(x-a)^N$$

de forma reducida

$$f(x) = \sum_{n=0}^{N} \frac{f^{n}(a)}{n!}(x-a)^n$$

Aproximaciones de la serie de Taylor

Si el valor de la función $f(x)$ es conocida en un punto x, entonces el valor de la función en su punto vecino $x + \Delta x$ puede ser estimada usando las series de Taylor. Si tomamos a $x = x_0 + h$, entonces $h = x - x_0$ y el polinomio quedaría como

$$f(x + \Delta x) = f(x) + \Delta x f'(x) + \frac{\Delta x^2}{2!} f''(x) + \frac{\Delta x^3}{3!} f'''(x) + \cdots$$

Reordenando

$$\frac{f(x + \Delta x) - f(x)}{\Delta x} = f'(x) + \frac{\Delta x}{2!} f''(x) + \frac{\Delta x^2}{3!} f'''(x) + \cdots$$

Aproximaciones de la serie de Taylor

Entonces la fórmula de derivada adelantada para evaluar la derivada de una función es

$$f'(x) = \frac{f(x + \Delta x) - f(x)}{\Delta x} + O(\Delta x)$$

donde la cantidad $O(\Delta x)$ representa que esta fórmula es de una precisión de primer orden.

Aproximaciones de la serie de Taylor

De manera similar la diferencia adelantada

$$f'(x) = \frac{f(x) - f(x - \Delta x)}{\Delta x} + O(\Delta x)$$

y la central

$$f'(x) = \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x} + O(\Delta x^2)$$

Esta ultima es de segundo orden.
La segunda derivada puede ser evaluada con

$$f''(x) = \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x)}{\Delta x^2}$$

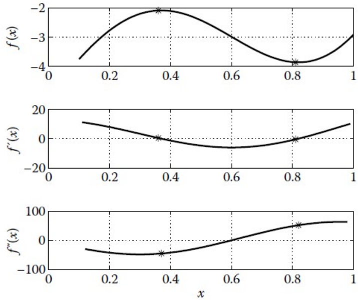
Ejercicio

Realice un código que estime la primera derivada (atrasada, adelantada y central) así como la segunda derivada. Una vez realizada la función calcule y grafique la función, la primera y segunda derivada de la siguiente función

$$f(x) = 2sin5x + 3x^3 - 2x^2 + 3x - 5$$

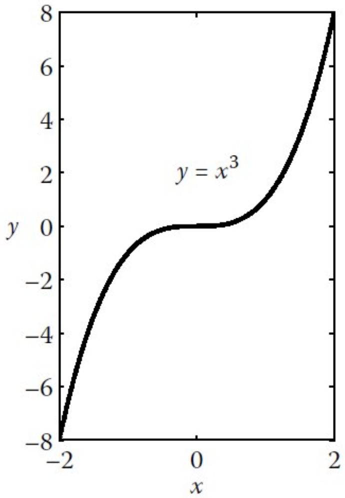
En el rango de 0 a 1 con un $\Delta x = 0.01$

Primera y segunda derivada



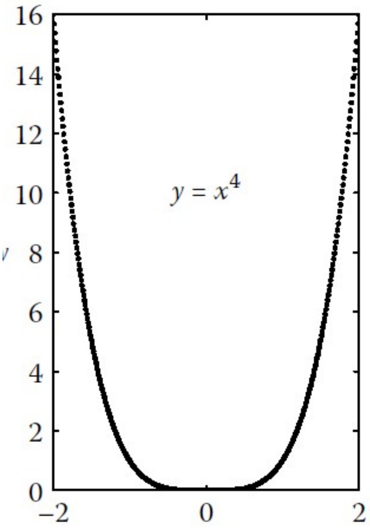
Note que la función tiene un máximo y un mínimo, se muestran con un *. De la primera derivada que $f'(x) = 0$ en el máximo y el mínimo. Ahora, la segunda derivada que $f''(x) \geq 0$ en el mínimo y es $f''(x) \leq 0$ en el máximo de la función. La segunda derivada proporciona información de curvatura de la función.

Primera y segunda derivada



Para lagunas funciones como $f(x) = x^3$, $f'(x) = f''(x) = 0$ en x=0; en tales casos, uno tiene que buscar derivadas de orden superior. Aquí $f'''(x) = 6$. La derivada de orden más alto que no sea cero se denota por n. Entonces x es un punto de inflexión si n es impar o es un optimo local si n es impar.

Primera y segunda derivada



Similar al caso anterior. Se puede observad que la función $f(x) = x^4$ tiene un mínimo local en x=0

Gradiente

El gradiente es un vector que contiene derivadas parciales de la función con respecto a la variables de diseño $(x_1, x_2, \ldots, x_n)$ y matemáticamente se expresa como:

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

Gradiente

La tangente y el gradiente, para el caso univarible son la mismo. Sin embargo, para el caso bivariable, la tangente para cada contorno de la función es diferente y el valor de la función permanece igual a lo largo de la tangente, es decir, a lo largo de una tangente

$$\nabla f = \frac{\partial f}{\partial x_1}\nabla x_1 + \frac{\partial f}{\partial x_2}\nabla x_2$$

Así el gradiente es normal a la tangente.

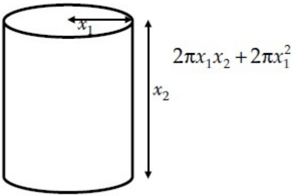
Gradiente

Propiedades importantes del vector gradiente

1. La magnitud del gradiente da la máxima tasa de incremente en $f(x)$.
2. El vector gradiente en el punto x is normal al hiperplano tangente definido por $f(x)$ constante.
3. La derivada direccional de $f(x)$ a lo largo de la dirección d, se define como $f_d'(x) = \nabla f(x)^T d = |\nabla f(x)||d|cos\theta$

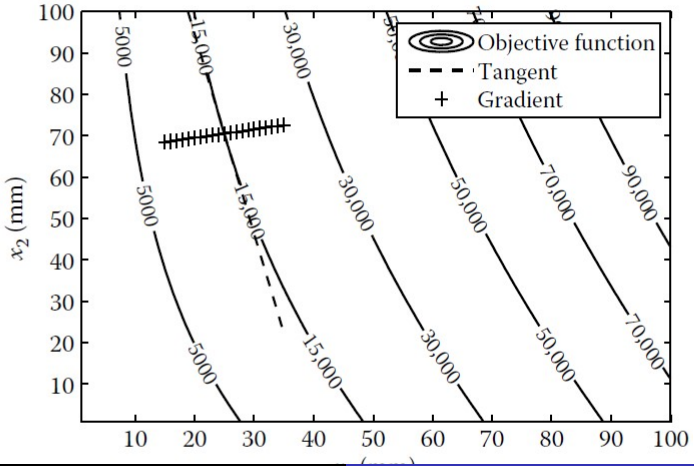
Gradiente

Ejemplo Un fabricante de refrescos debe producir una lata cilíndrica que pueda contener 330 ml de un refresco. Quiere hacer las dimensiones del contenedor de tal manera que se minimice la cantidad de material utilizado en su construcción. Las variables de diseño para el problema de optimización son el radio y la altura de la lata.La lata cilíndrica consiste en una porción curva y dos extremos circulares. El área de la porción curva está dada por $2\pi x_1 x_2$ y el area de las tapas circulas esta dada por $2\pi x_1^2$



Gradiente

Ejemplo.



Gradiente

Ahora si consideramos tres funciones $f_1(x_1, x_2, x_3)$, $f_2(x_1, x_2, x_3)$ y $f_3(x_1, x_2, x_3)$, las cuales están en función de tres variables x_1, x_2, x_3 el gradiente de estas funciones puede puesta en un matriz simple llamada Jaccobiana J y se expresa en forma matemática como

$$[J] = \begin{bmatrix} \frac{\partial f_1}{dx_1} & \frac{\partial f_1}{dx_2} & \frac{\partial f_1}{dx_3} \\ \frac{\partial f_2}{dx_1} & \frac{\partial f_2}{dx_2} & \frac{\partial f_2}{dx_3} \\ \frac{\partial f_3}{dx_1} & \frac{\partial f_3}{dx_2} & \frac{\partial f_3}{dx_3} \end{bmatrix}$$

Derivada direccional

En optimización, es posible que moverse en la dirección del gradiente puede resultar en una región no factible. En tal caso, uno desea moverse en alguna otra dirección y le gustaría saber la velocidad de cambio de la función en esa dirección. La derivada direccional proporciona información sobre la tasa instantánea de cambio de una función en una dirección particular. Si $\mathbf{u}$ es un vector unitario, entonces la derivada direccional de la fución $f(x)$ en dirección de $\mathbf{u}$, esta dada por

$$\nabla f(x)^T \mathbf{u}$$

Matriz Hessiana

La matriz Hessiana $\mathbf{H}$ representa la segunda derivada de una función con más de una variable. Para una función $f(x_1, x_2, x_3)$ con tres variables, la matriz Hessiana se escribe como

$$[\mathbf{H}] = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & \frac{\partial^2 f}{\partial x_1 x_3} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 x_3} \\ \frac{\partial^2 f}{\partial x_3 x_1} & \frac{\partial^2 f}{\partial x_3 x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix}$$

Matriz Hessiana

La matriz Hessiana debe ser definida positiva en el mínimo de la función. Una matriz es definida positiva si sus eigenvalores son positivos. Para una matriz cuadrada, existe un vector no cero tal que cuando es multiplicado con la matriz cuadrada este resulta en un vector que difiere del original por un multiplicador escalar. El vector no cero esta en terminos del eigenvector y el escalar multiplicador eigenvalor.

Aproximaciones a la segunda derivada parcial

$$\frac{\partial f}{\partial x_i} = \frac{f(x_i + \Delta x_i) - f(x_i - \Delta x_i)}{2\Delta x_i}$$

$$\frac{\partial^2 f}{\partial x_i^2} = \frac{f(x_i + \Delta x_i) - 2f(x_i) + f(x_i - \Delta x_i)}{\Delta x_i^2}$$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \Big[f(x_i + \Delta x_i, x_j + \Delta x_j) - f(x_i + \Delta x_i, x_j - \Delta x_j) - f(x_i - \Delta x_i, x_j + \Delta x_j) + f(x_i - \Delta x_i, x_j - \Delta x_j) \Big] \Big/ (4\Delta x_i \Delta x_j)$$

Matriz Hessiana

Ejercicio. Encuentre el gradiente y la matriz Hessiana de la función

$$f(x) = x_1^2 + 2x_1x_2 + 3x_2^2 + 4x_3^2 - 5x_2x_3$$

También encuentre la derivada direccional de la función en el punto (1,1,1) en la dirección

$$d = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Aproximación lineal y cuadrática

Usando la serie de Taylor, encuentre la aproximación lineal y cuadrática de

$$f(x) = 3x_2 - \frac{x_1}{x_2}$$

en el punto $(2, 1)$

Aproximación lineal y cuadrática

Usando la serie de Taylor, encuentre la aproximación lineal y cuadrática de

$$f(x) = 2\pi x_1 x_2 + 2\pi x_1^2$$

en el punto $(80, 100)$