Tema 1 : Optimizadores

En este tema se presenta la idea básica de la optimización de funciones, la cual consiste idealmente en encontrar la ubicación del mínimo global, es decir, para una función $y = f(x)$, encontrar el valor de x que produzca el menor valor para y y que $dy/dx = 0$. Esto puede llevarse a cabo de manera analítica o numérica. En la forma analítica es necesario obtener la derivada de $y = f(x)$ e igualar a cero la solución, es decir $dy/dx = df(x)/dx = 0$ y posteriormente resolver para x.

Como ejemplo se tiene la función $f(x) \in \mathbb{R}^1$ mostrada en la Figura 1 que, al poder ser representada en un grafica de dos dimensiones, se puede apreciar la ubicación del mínimo que es en $x = 5$. Analíticamente, podemos calcular la derivada como:

$$\frac{dy}{dx} = \frac{d\{(x-5)^2 + 10\}}{dx} = 2(x-5) = 0 \tag{1}$$

por tanto, resolviendo para x tenemos que $x = 5$.

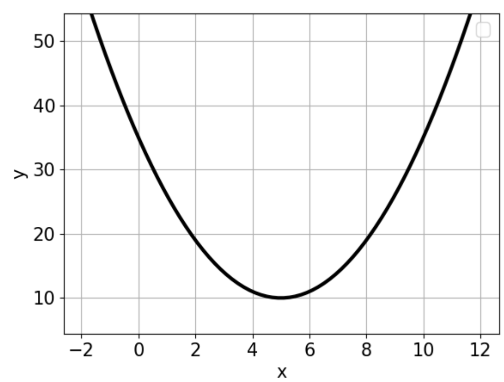


Figure 1: Gráfica de la función $y = f(x) = (x-5)^2 + 10$.

Sin embargo, en aplicaciones mas avanzadas o escenarios reales es común tener que optimizar funciones mas complejas $y = f(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ y con un gran número de variables de entrada x, por lo que optimizar de manera analítica suele resultar ser inviable. Por esta razón se ha optado por utilizar el método numérico o generalmente llamado *optimización a pasos*.

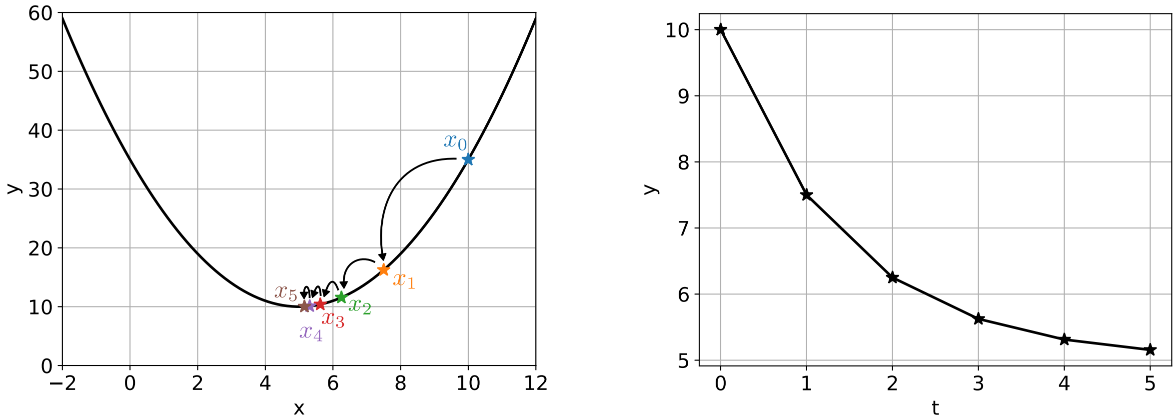
Optimización a pasos (Gradiente descendente)

Con el método de optimización a pasos es posible encontrar un mínimo (Puede ser global o local) de una función $f(x) \in \mathbb{R}^n$ arrojando una solución x^* *aproximada* al problema de minimización:

$$f(x^*) = \min_{x \in \mathbb{R}^n} f(x) \tag{2}$$

Este proceso sigue de forma general el siguiente procedimiento:

1. Asignar un valor inicial para x_0 y evaluar $y = f(x_0)$.
2. Determinar si x_0 debe incrementarse o reducirse afín de obtener un menor valor de y.
3. Aumentar o reducir (Dependiendo del paso 2) el valor de x_0
4. Evaluar $y = f(x_0)$.
5. Repetir los pasos 2,3 y 4 incrementando en uno el subíndice de x



(a) (b)

Figure 2: a) Función $y = f(x) = (x - 5)^2 + 10$ donde se ilustra el valor de x inicial (x_0) así como los valores de x obtenidos en cada iteración t b) Valores de $y = f(x)$ para cada valore de x_t ($f(x_t)$)

El objetivo de este proceso es encontrar el menor valor $f(x)$ para un valor de x tal que $df(x)/dx = 0$. Para determinar si x_0 debe incrementarse o reducirse (Paso 2) es posible emplear el signo del resultado de la derivada $df(x_t)/dx$ (Pendiente) donde x_t es el valor de x en la iteración t. Asimismo, el nuevo valor de x se aumentará o reducirá proporcionalmente al valor de $df(x_t)/dx$ resultando así la siguiente *regla de actualización* (Paso 3):

$$x_t = x_{t-1} - \alpha \nabla f(x_{t-1}) \tag{3}$$

Donde típicamente $\alpha \in (0, 1]$.

El método de optimización a pasos en general permite reducir el valor de $f(x)$ encontrando un valor de x que se aproxima a la ubicación de un mínimo local

Ejemplo resuelto

Tómese como ejemplo la función mostrada en la Figura 1 y considere un domino para la variable independiente x de $[-2, 12]$. Utilizando el método de gradiente descendente (GD) se buscará el valor de x que produzca el menor valor de $f(x)$ cuya derivada en x sea cero. Primero se define arbitrariamente $x_0 = 10$ (Paso 1). Después se evalúa el valor de la derivada en este punto, el signo de esta derivada determina si x_0 debe incrementarse o reducirse y ser asignado a x_1 (Paso 2). Posteriormente se actualiza el valor de x ($x_0 \rightarrow x_1$) de acuerdo a la Ecuación 3 (Paso 3). Con el nuevo valor de x se deberá calcular el valor de $y = f(x)$ afín de verificar que su valor este disminuyendo (Paso 4). Este proceso (desde el paso 2 hasta el 4) deberá repetirse t iteraciones o hasta que el valor de x_t ya no este cambiando sustancialemtne de iteración a iteración.

En la Figura 2a se ilustra el valor de x_t para 5 iteraciones del procedimiento de optimización. El la Figura 2b se muestra el valor de $y = f(x)$ obtenido para cada valor de x_t. Específicamente, esta Figura 2b es de utilidad para evaluar el funcionamiento correcto del algoritmo de optimización ya que se aprecia que el valor de $y = f(x)$ disminuye en cada iteración.

Adicionalmente se puede verificar el valor exacto de x_t y el valor de la derivada dy/dx para cada x_t obtenido. Esta información se muestra en la Tabla 1.

	$t=0$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
x_t	10	7.5	6.25	5.625	5.3125	5.1562
dy/dx	10	5	2.5	1.25	0.625	0.3125

Table 1: Valores de dy/dx para cada valor obtenido de x_t.

Nótese que el valor α debe ser cuidadosamente seccionado afín de que el algoritmo de optimización tenga éxito; Un valor muy pequeño de α ($\alpha \approx 0$) implicará que demasiadas iteraciones sean necesarias para poder encontrar el valor x donde se encuentra el mínimo. Por otro lado un valor alto de α causará inestabilidad en el proceso teniendo cada nuevo valor de x mas alejado del mínimo. Para este ejemplo se usó $\alpha = 0.25$.

En la Tabla 2 se muestra la comparación entre los dos métodos de optimización; El método analítico y el numérico.

Comparación entre métodos			
Método	α	Iteraciones	Ubicacion en x del Mínimo
Analítico	-	-	5
Numérico	0.25	5	5.1562

Table 2: Comparación del mínimo alcanzado con cada método.

Ejercicios

1. Proponer una función $y = f(x)$ (por ejemplo $y = x^2$) cuya característica principal sea que tenga un mínimo global. A partir de ésta, implementar un algoritmo numérico basado en gradiente descendente (GD) para encontrar el valor para x donde se encuentra ese mínimo. Se deberá hacer una comparación del método numérico contra el método analítico.

2. Proponer una función $y = f(x) + g(x) + \ldots$ que sea el resultado de la suma de al menos otras dos funciones (por ejemplo $g(x) = (x-3)^2 + 8$ y $f(x) = 20sen((x+\pi)*5) - 1$). La idea es tener como referencia una función $y = f(x) + g(x) + \ldots$ que tenga al menos dos mínimos, preferentemente cada mínimo en distinta posición en y. Posteriormente, implementar un algoritmo numérico basado en gradiente descendente (GD) para encontrar la posición en x de un mínimo. Se deberá hacer una comparación del método numérico contra el método analítico.