1 Optimizacion sin restriccion

Considere el siguiente problema

$$\begin{matrix} Minf(x) \\ x \in \Omega \end{matrix} \tag{1}$$

Donde Ω generalmente es todo R^n o un subconjunto sencillo de R^n. A continuacion describiremos las propiedades que tiene que tener el punto que sea solucion del problema.

1.1 Condicion necesaria de primer orden

Sea x_0 un punto $\in R^n$ donde ocurre el minimo de la funcion. Sea x un punto cercano a x_0 tal que $x = x_0 + h$ donde $h \in R^n$ y esta en cualquier direccion. La aproximacion en series de Taylor de primer orden en el punto x_0 sera.

$$f(x_0 + h) = f(x_0) + \nabla f(x_0)h + R(h) \tag{2}$$

Planteando que $h = t\Delta x$. Donde Δx es un vector unitario en cualquier direccion y "t" es escalar que indica el largo del vector unitario. Podemos plantaer entonces

$$f(x_0 + t\Delta x) = f(x_0) + t\nabla f(x_0)\Delta x + R(h) \tag{3}$$

Dado que suponemos que $f(x_0)$ es un minimo, debe cumplirse que

$$f(x_0) \leq f(x_0 + t\Delta x) \tag{4}$$

Etonces si dividimos por "t" y lo hacemos tender a cero se obtiene.

$$0 \leq \nabla f(x_0)\Delta x \tag{5}$$

Dado que consideramos cualquier direccion posible se establece que la condicion necesaria de primer orden es

$$\nabla f(x_0) = 0 \tag{6}$$

1.2 Condicion necesaria de segundo orden

Para derivar condiciones necesarias de segundo orden debemos realizar la aproximacion en series de taylor en el punto x_0 hasta la segunda derivada

$$f(x_0 + t\Delta x) = f(x_0) + t\nabla f(x_0)\Delta x + \frac{1}{2}t\Delta^T x H(x_0)\Delta x + R(h) \qquad (7)$$

Considerando la ecuacion 5 y 6 se deriva que si x_0 es un minimo debe cumplirse que

$$0 \leq \Delta^T x H(x_0)\Delta x \qquad (8)$$

1.3 Condiciones suficientes

Notemos que las condiciones que se derivaron en los apartados anteriores son solo condiciones necesarias pues se partio del echo de que x_0 era un punto extremo (maximo o minimo local). Para obtener una condicion suficiente debemos partir sin suponer nada de la naturaleza del punto. De este modo para obtener una condicion suficiente: Sea la aproximacion en series de taylor en un punto x_0

$$f(x_0 + t\Delta x) = f(x_0) + t\nabla f(x_0)\Delta x + \frac{1}{2}t\Delta^T x H(x_0)\Delta x + R(h) \qquad (9)$$

Supongase que

$$\begin{aligned} \nabla f(x_0) &= 0 \\ 0 &\leq \Delta^T x H(x_0)\Delta x \end{aligned} \qquad (10)$$

reemplazando en la ecuacion 9 se tiene que x_0 tiene que ser un minimo local

2 Conjuntos convexos

Sea $U = u_1, u_2, ..., u_m$ conjunto formado por elementos de R^n. Una combinacion convexa es un una conbinacion de los elementos de U tal que

$$u_{m+1} = \alpha_1 u_1 + \alpha_2 u_2 + ... + \alpha_m u_m \qquad (11)$$

con la condicion de que

$$\alpha_i \geq 0$$
$$\sum_{i=1}^{n} \alpha_i = 1$$

Un conjunto convexo es un subconjunto $C \in R^n$ tal que para todo par de puntos en C (u_1, u_2), cualquier conbinacion convexa genera un elemento que tambien pertenece a "C"

3 funciones convexas

1. Sea un conjunto $X \in R^n$ convexo. Una funcion definida en "X" es una funcion convexa si para dos puntos cualesquiera en "X" (x_1, x_2) y para todo $\lambda \in R$ con $0 \leq \lambda \leq 1$ ocurre que

$$f[\lambda x_2 + (1-\lambda)x_1] \leq \lambda f[x_1] + (1-\lambda)f[x_2] \tag{12}$$

2. Sea un conjunto $X \in R^n$ convexo. Una funcion definida en "X" es una funcion concava si para dos puntos cualesquiera en "X" (x_1, x_2) y para todo $\lambda \in R$ con $0 \leq \lambda \leq 1$ ocurre que

$$f[\lambda x_2 + (1-\lambda)x_1] \geq \lambda f[x_1] + (1-\lambda)f[x_2] \tag{13}$$

3. Sea f_1 y f_2 funciones convexas sobre un conjunto convexo Ω entonces la funcion $g = f_1 + f_2$ es una funcion convexa sobre Ω

4. Sea "f" una funcion convexa sobre un conjunto convexo Ω. entonces la funcion a*f, es una funcion convexa para todo $a \geq 0$

5. Sea f un funcion convexa sobre un conjunto convexo Ω. El conjunto $A = x : x \in \Omega, f(x) \leq c$ es convexo para todo numero real c

6. Si f $\in C^1$ .entonces f es convexa sobre un conjunto convexo Ω si y solo si

$$f(x_2) \geq f(x_1) + \nabla f(x_1)(x_2 - x_1) \tag{14}$$

para todo $x_1, x2 \in \Omega$ Aveces conbiene ver la expresion 14 de otra manera. Considerando que $h = x_2 - x_1$ la ecuacion 14 queda

$$f(x_1 + h) \geq f(x_1) + \nabla f(x_1)h \tag{15}$$

7. Sea f(x) $\in C^2$. entonces f(x) es una funcion convexa sobre un conjunto convexo conteniendo un punto interior si y solo si la matriz hesiana de f(x) es definida semipositiva

4 extremos de funciones convexas

1. Sea f(x) funcion convexa definida en un conjunto convexo Ω. Sea "P" el conjunto de los puntos donde f(x) tiene sus minimos locales. etonces "P" es un conjunto convexo. Ademas todo minimo local es ademas un minimo global.

   (a) Primero demostraremos el echo que todo minimo local es un minimo global Sea f(x) una funcion convexa. Sea x_0 un minimo relativo y Sea x_* el minimo global. y concedase que

   $$f(x_0) > f(x_*) \tag{16}$$

   Dado que f(x) es convexa entonces para todo λ que cumpla $0 \leq \lambda \leq 1$ se tiene

   $$f[\lambda x_* + (1-\lambda)x_0] \leq \lambda f(x_*) + (1-\lambda)f(x_0) \tag{17}$$

   Reemplazando la ecuacion 16 en la 17 se tiene que

   $$f[\lambda x_* + (1-\lambda)x_0] < \lambda f(x_0) + (1-\lambda)f(x_0) = f(x_0) \tag{18}$$

   Es decir siempre puedo escojer un λ que me permita obtener un valor $x = \lambda x_* + (1-\lambda)x_0$ en las vecindades de x_0 para que el se puede establecer que $f(x) < f(x_0)$ lo que contradice el hecho de que x_0 es un minimo relativo

   (b) Probemos ahora el echo de que los puntos donde ocurren los minimos (que son globales) forman un conjunto convexo, que de otra manera dicen que ocurren todos juntos.

   Supongase que el minimo global es tomado en dos puntos dintintos (no supongamos nada sobre su cercania aun).Sean x_1, x_2los puntos aquellos. Dado que la funcion es convexa se debe cumplir que

   $$f(\lambda x_2 + (1-\lambda)x_2) \leq \lambda f(x_2) + (1-\lambda)f(x_1) \tag{19}$$

   dado que $f(x_1)$ y $f(x_2)$ son minimos globales (es decir tienen el mismo valor)la ecuacion 19 queda

   $$f(\lambda x_2 + (1-\lambda)x_2) \leq \lambda f(x_2) + (1-\lambda)f(x_2) = f(x_2) \tag{20}$$

Sea $x = \lambda x_2 + (1-\lambda)x_2$ entonces la ecuacion 20 queda finalmente

$$f(x) \leq f(x_2) \tag{21}$$

Dado que $f(x_2)$ es el minimo global el menor valor que puede tomar f(x) es precisamente $f(x_2)$, es decir $f(x) = f(x_2)$. Lo que hemos probado es que si hay mas de un punto donde se alcanza el minimo global, todos los puntos que se generen por combinacion convexa de ellos tambien seran minimos globales. Dado que estos puntos perteneces a un conjjunto convexo. Los puntos donde se alcanza el minimo forman un conjunto convexo

2. Ahora sea f(x) un funcion convexa sobre una region convexa, sea ademas $f(x) \in C^1$. Sea x_0 un punto done $\nabla f(x_0 = 0)$. Por la ecuacion 14 tenemos

$$\begin{gathered} f(x_2) \geq f(x_0) + \nabla f(x_0)(x_2 - x_1) \\ f(x_2) \geq f(x_0) \end{gathered} \tag{22}$$

Es decir x_0 es un minimo local y por consiguiente un minimo global Es decir en el caso de funciones convexas el criterio de la primera derivada es razon suficiente para asegurar que existe un minimo

5 algoritmos optimizacion no restringida

5.1 Problemas unidimensinales no restringidos

5.1.1 Metodo de Fibonacci y la seccion aurea

Una funcion es unimodal en un intervalo si solo tiene un maximo (o minimo) relativo en el intervalo. Como es de esperar la funciones convexas son ejmeplos de funciones unimodales. El concepto de unimodal incluye a funciones no diferenciables y funciones que no sean continuas. Considere entonces que el problema consiste en encontrar el minimo de una funcion unimodal $f(x)$ El razonamiento del algoritmo es el siguiente:

- Sea que se busca el minimo de la funcion unimodal $f(x)$ en el intervalo cerrado [a,b]

- Sea que se conoce el valor de la funcion (no necesariamente la funcion) en puntos interiores del intervalo

$$a < x_1 < x_2 < x_3 < ..... < x_n < b$$

- Sea que se evalua la funcion en dos puntos interiores x_k y x_l de manera que sucede $f(x_k) < f(x_l)$. Dado que a funcion es unimodal podemos asegurar que el minimo esta en el intervalo $[a, x_k]$.

- El metodo de fibonacci busca mediante un procedimieto iterativo reducir el intervalo donde buscar el minimo. A continuacion se muestra los pasos del algoritmo

  1. Se busca el minimo de la funcion unimodal $f(x)$ en intervalo [a,b].Supongase N =4 un entero elejido. El intervalo se escoje tras realizar el siguiente calculo

  $$\begin{aligned} z_0 = a_0 + \alpha_0[b_0 - a_0] \\ y_0 = b_0 - \alpha_0[b_{0-a_0}] \end{aligned} \tag{23}$$

  Donde

  $$\alpha_0 = \frac{F_{N-i-1}}{F_{N-i}} = \frac{F_{4-0-1}}{F_{4-0}} = \frac{F_3}{F_4} \tag{24}$$

  Donde F_j corresponde al j-esimo termino de la serie de fibonacci. En nuestro caso la ecuacion ”n” quedaria

  $$\begin{aligned} z_0 = a_0 + \tfrac{3}{5}[b_0 - a_0] \\ y_0 = b_0 - \tfrac{3}{5}[b_{0-a_0}] \end{aligned} \tag{25}$$

  2. Se compara el valor de $f(z_0)$ con el de $f(y_0)$

  $$Sif(y_0) < f(z_0)$$

  $$a_1 = a_0$$

  $$b_1 = z_0$$

  $$Sif(y_0) > f(z_0)$$

  $$a_1 = y_0$$

  $$b_1 = b$$

3. ahora el intervalo a considerar (supongase $f(y_0) < f(z_0)$) es $[a_0, z_0]$. Primero es calcular α_1

$$\alpha_1 = \frac{F_{N-i-1}}{F_{N-i}} = \frac{F_{4-1-1}}{F_{4-1}} = \frac{F_2}{F_3} = \frac{2}{3} \tag{26}$$

   Ahora la ecuacion "n" quedara

$$\begin{array}{l} z_1 = a_1 + \alpha_1[b_1 - a_1] = a_0 + \frac{2}{3}[z_0 - a_0] \\ y_1 = b_1 - \alpha_1[b_1 - a_1] = z_0 - \frac{2}{3}[z_0 - a_0] \end{array} \tag{27}$$

4. Se vuelve al paso "n" para obtener un nuevo intervalo
5. El ultimo paso de la iteracion lleva al mismo punto en ambos, por lo que no es posible seguir con la iteracion

Veamos un ejemplo

$$MINf(x) = 2 + |x - 1|$$

$$0 \leq x \leq 3$$