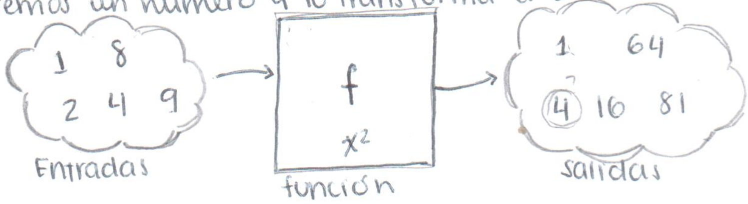
Ejercicio 1:

Cuando buscamos minimizar una función f, buscamos encontrar la entrada x^* que nos devuelva el menor valor de $f(x)$.

Esto es, tenemos una maquina f, en dónde le metemos un número y lo transforma a otro



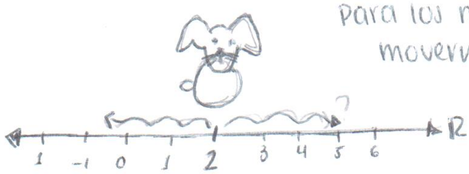
¿como encontramos la entrada que nos da el menor valor sin tener que revisar todas las entradas?

una opción es el algoritmo de búsqueda lineal

En este algoritmo primero elegimos una entrada inicial y

$2 \to \boxed{f} \to 4$

Ahora elegimos una dirección en la que nos vamos a mover



Para los números en la recta real, podemos movernos a la izquierda (-1) o a la derecha (+1)

(si decidimos movernos a la derecha nuestra función va incrementar, por lo que es una mala dirección)

si primero decidimos movernos a la izquierda, ahora queremos ~~dar~~ ver ~~cuántos pasos~~ que tan grande debe ser nuestro salto a la izquierda.

Si nuestro salto es muy pequeño

$1.99 \to \boxed{f} \to 3.96$ nuestra función no ha decrecido mucho y tardaremos mucho en llegar a una solución

Si nuestro salto es muy grande

$-3 \to \boxed{f} \to \boxed{9}$ nuestra función puede crecer, que es lo que no queremos.

Por eso buscamos el salto que más nos acerca al mínimo

es decir $\min_{\alpha > 0} f(x + \alpha p)$

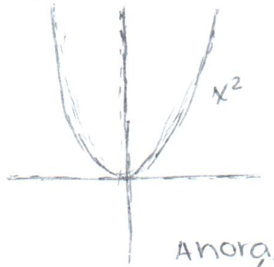
En nuestro ejemplo este sería $\alpha = 2$ ya que

$f(x+\alpha p) = f(2+2(-1)) = f(0) = 0$

Que es el valor mínimo que puede tomar un cuadrado

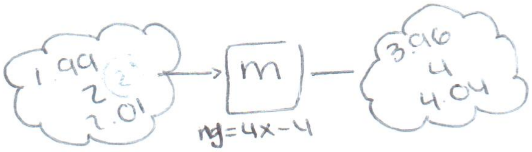
Así, vimos que para este método es importante la dirección que tomes como el tamaño del salto que das.

Región de confianza



para este algoritmo, sabemos que nuestra maquina nos devuelve la función x^2, pero esta es muy dificil calcular, entonces tomamos un modelo que se parece al nuestro, pero es más fácil calcular

Ahora, si de nuevo elegimos nuestra x inicial como 2 nosotros tenemos otra función m que se parece mucho a f para valores cercanos a 2



Estos valores son muy cercanos a los verdaderos (3.9601, 4, 4.0401)

Ahora, este modelito nos queda $m(x+p)=4(x+p)-4$

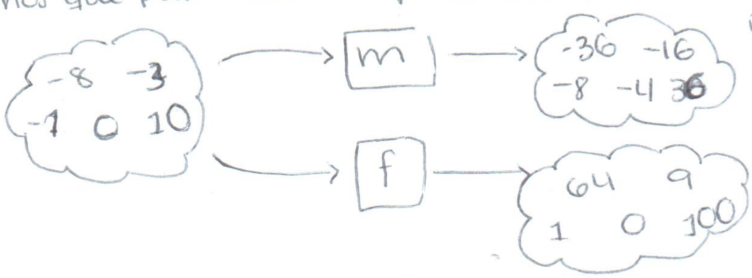
Que dado que $x=2$ queda $m(2+p)=4p-8$

Esto es más facil de minimizar, pues si aumentamos p crece m y si disminuimos p disminuye f.

Entonces podríamos que un p muy negativo va ser la respuesta, pero esto es incorrecto.

¿Por qué?

veamos que pasa cuando p decrece mucho



como vemos para valores muy diferentes de 2

nuestro modelo ya no da valores cercanos a los reales

Por eso, en este modelo generamos una región de confianza, que es una región en dónde se sabe que el modelo aproxima bien a la función

Por eso, cuando utilizamos el algoritmo de región de confianza es importante la región de confianza que se elije para el punto.

Ejercicio 2:

Sea $f(x)=\frac{1}{2}x^TQx - b^Tx$, buscamos el minimizador de la $\phi(\cdot)$ definido como

$$\phi(\alpha) = f(x_k + \alpha p_k)$$

Ahora, cualquier minimizador α de $\phi(\alpha)$ satisface que $\phi'(\alpha)=0$

Así, sea α^* minimizador de $\phi(\alpha)$ tenemos que

$$(1)\ \phi'(\alpha^*) = \nabla f(x_k + \alpha^* p_k)^T p_k = 0$$

Luego, tenemos que $\nabla f(x) = Qx - b$

Así, por la ecuación (1) obtenemos

$$[Q(x_k + \alpha^* p_k) - b]^T p_k = 0$$

$$\Rightarrow [Qx_k - b + \alpha^* Q^T p_k]^T p_k = 0$$

$$\Rightarrow [Qx_k - b]^T p_k + \alpha^* p_k^T Q p_k = 0$$

Notamos que al ser f convexa tenemos que existe un único mínimo

Notando que $\nabla f(x) = Qx - b$ y sustituyendo tenemos

$$\nabla f(x_k)^T p_k + \alpha^* p_k^T Q p_k = 0$$

Por último, despejando obtenemos

$$\alpha^* = \frac{-\nabla f(x_k)^T p_k}{p_k^T Q p_k} = \frac{-\nabla f_k^T p_k}{p_k^T Q p_k}$$