Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ingeniería Industrial y de Sistemas
ICS1102 Optimización – 2do Semestre 2008 – secciones 2,3 y 4

PROFESORES: Juan Carlos Muñóz, Pablo Rey ,Sergio Toloza
AYUDANTE: Mathias Klapp

Ayudantía N°5: Métodos Numéricos de Optimización

OPTIMIZACIÓN NUMÉRICA SIN RESTRICCIONES:

a) MÉTODO DE NEWTON UNIDIMENSIONAL

Funciona aproximando la F.O. por parábolas y se deduce del polinomio de Taylor:

$$f(x) \approx f(x_k) + f'(x_k)(x - x_k) + \frac{f''(x_k)(x - x_k)^2}{2}\Bigg| \frac{d}{dx} \quad \rightarrow \quad f'(x) \approx f'(x_k) + f''(x_k)(x - x_k) = 0$$

$$\boxed{x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}}$$

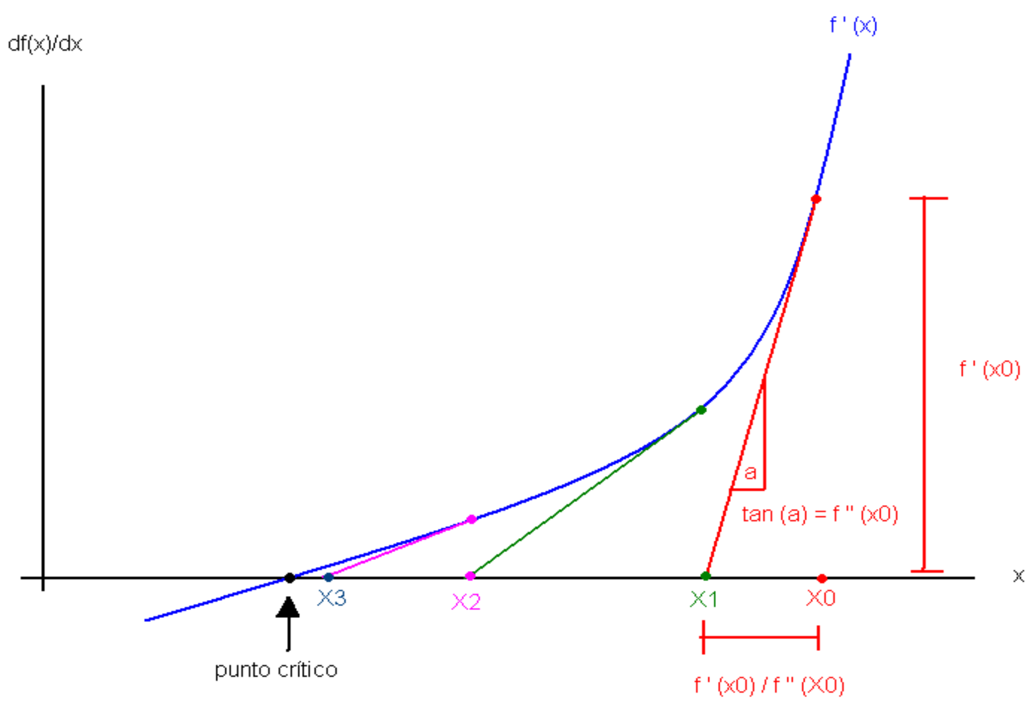
OJO:

→ Caso útil: Cuando despejar x^* de $\frac{df(x^*)}{dx} = 0$ no es trivial

→Solo funciona con una dimensión y no siempre converge.

→Sólo detecta puntos críticos, no distingue Max con Min local (revisar $D^2 f(x)$).

→Se detiene cuando: $f(x_k) - f(x_{k+1}) \longrightarrow 0$



Problema 1 (Resolución numérica por Newton):

Resolver $Min: x^4 - 26x^2 + 48x + 10$ a través del método de Newton partiendo de los puntos $x_0 = 0; 5$ y -5. ¿Puede decidir con sus resultados cual es la solución óptima?

Respuesta:

Se sabe que: $f`(x) = 4x^3 - 52x + 48$ y que $f``(x) = 12x^2 - 52$.

Con $X_0 = 0$ →

$$x_1 = x_0 - \frac{f`(x_0)}{f``(x_0)} = 0 + \frac{48}{52} = 0.923$$

$$x_2 = x_1 - \frac{f`(x_1)}{f``(x_1)} = 0.923 + 0.075 = 0.998$$

$$x_3 = x_2 - \frac{f`(x_2)}{f``(x_2)} = 0.998 + 0.001 = 0.999 \approx 1$$

Como sabemos que f``(1) = - 40 < 0 , este punto no nos sirve ya que es un máximo local.

Con $X_0 = 5$ →

$$x_1 = x_0 - \frac{f`(x_0)}{f``(x_0)} = 5 - \frac{f`(5)}{f``(5)} = 3.839$$

$$x_2 = x_1 - \frac{f`(x_1)}{f``(x_1)} = 3.839 - \frac{f`(3.839)}{f``(3.839)} = 3.355$$

$$x_3 = x_2 - \frac{f`(x_2)}{f``(x_2)} = 3.355 - \frac{f`(3.355)}{f``(3.355)} = 3.059$$

$$x_4 = x_3 - \frac{f`(x_3)}{f``(x_3)} = 3.059 - \frac{f`(3.059)}{f``(3.059)} = 3.002 \approx 3$$

Luego f``(3) = 56 > 0, sabemos que es un mínimo local.

Con $X_0 = -5$ →

$$x_1 = x_0 - \frac{f`(x_0)}{f``(x_0)} = -5 - \frac{f`(-5)}{f``(-5)} = -4.226$$

$$x_2 = x_1 - \frac{f`(x_1)}{f``(x_1)} = -4.226 - \frac{f`(4.226)}{f``(4.226)} = -4.016$$

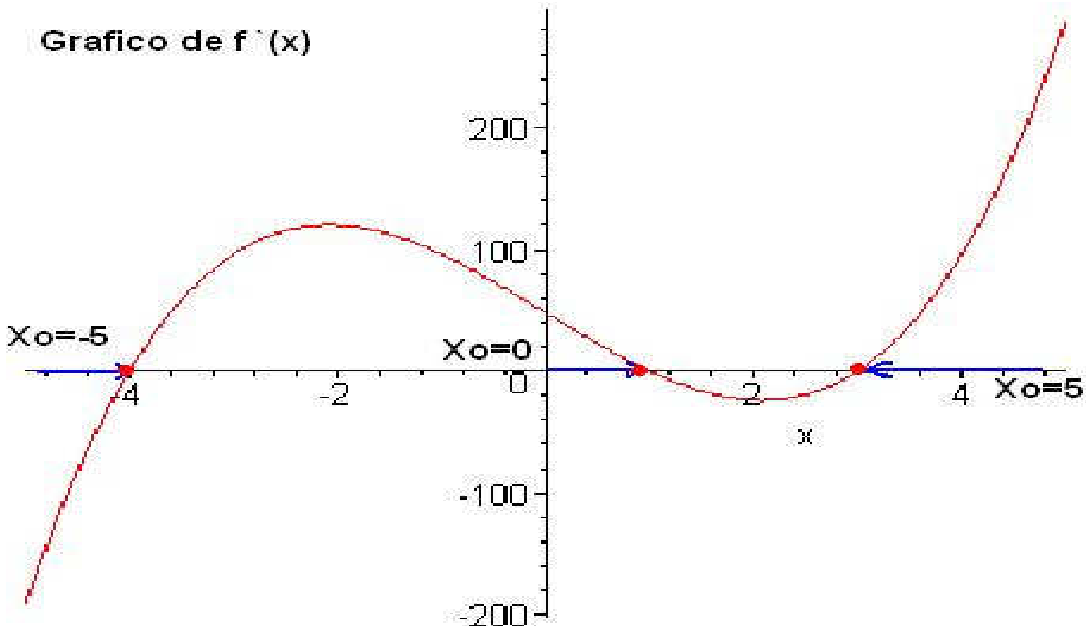
$$x_3 = x_2 - \frac{f`(x_2)}{f``(x_2)} = -4.056 - \frac{f`(-4.056)}{f``(-4.056)} = -4.001 \approx -4$$

Luego f``(- 4) = 140 > 0, sabemos que es un mínimo local.

No diverge en los bordes, por lo que podemos afirmar que existe solución.

Comparando f(3) = -467 < f(-4) = 1738.

→ Solución Optima es x = 3.



Problema 2 (comente):
Verdadero o falso.
"Si bien el método de newton no garantiza convergencia, cualquier polinomio de grado 2 es resuelto por el método en una iteración"

Respuesta:
Verdadero. El método converge en una iteración debido a que estamos aproximando a una parábola por parábolas, es decir, lo que realmente se está haciendo es resolver el problema original analíticamente. Tarea Propuesta: Ver esto con un ejemplo....

b) MÉTODO DEL GRADIENTE:

→ Parte de un punto X_0 y busca la dirección de máximo decrecimiento (caso Min) recorriendo la función hasta dejar de decrecer.

→ Garantiza convergencia a un mínimo local (¡No global!). En la práctica transforma un problema de "n" dimensiones a varios iterativos unidimensionales[1].

→ Algoritmo:

Paso Inicial:

- Se elige un vector inicial X_0 y se calcula $\nabla f(X)$.

Paso Iterativo:

- Se construye $X_{k+1} = X_k - \nabla f(X_k) * d$, en donde $-\nabla f(X_k)$ es la dirección de máximo decrecimiento y con $d \geq 0$[2].
- Se encuentra el 'd' óptimo del problema unidimensional:

$$\text{Min } f(X_{k+1}(d))$$

- Con d se construye X_{k+1}

Término:

- Se prueba el criterio de término (por ejem: $X_{k+1} - X_k \longrightarrow 0$)

---

[1] Los que se pueden resolver por Newton unidimensional.....

[2] Esta restricción implica que hacemos decrecer la función. Es decir multiplicamos al vector de decrecimiento por algo positivo y se lo sumamos al próximo punto.

Problema 3:

a) Dibuje aproximadamente, en esta misma hoja, el recorrido de **descenso** que haría una persona que camine de 1 a 2, siguiendo las reglas del método del Cauchy o Gradiente.

b) ¿qué ángulo forman entre sí cada etapa del camino que recorrerá la persona?

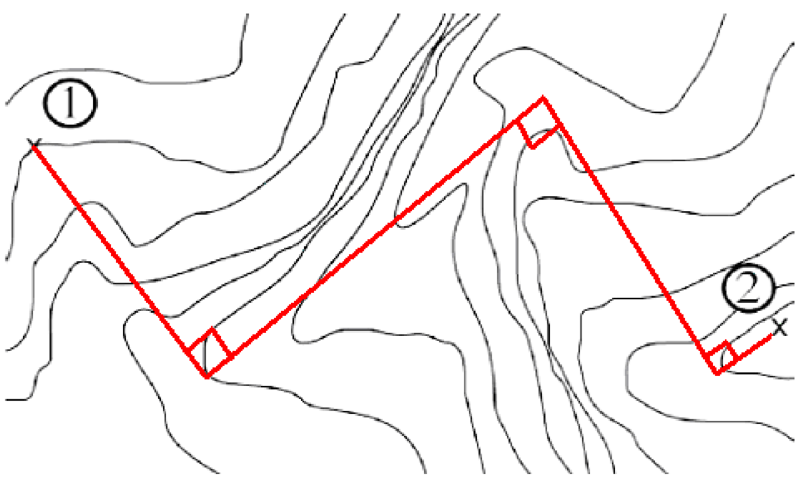
c) En cualquier punto (x,y) intermedio de una etapa del recorrido, el valor de

$$-\vec{\nabla} f(x, y) \cdot \vec{d}$$

donde d es la dirección hacia la cual camina la persona, es ¿ siempre positivo, negativo, o puede cambiar de signo en la etapa? Justifique.

Respuesta:

a)



b) Forman 90 grados.

c) d es siempre positivo, de lo contrario se estaría yendo por una dirección de máximo crecimiento con ‘d’ negativo.

Problema 4:

Resolver a través del método del Gradiente: Min $\frac{x^2}{4}+y^2$.

Parta primero con el punto (2,1). ¿Cuál es el óptimo? Pruebe luego partiendo del (1.0) y explique la diferencia en la convergencia del método.

Respuesta:

Sabemos que: $\nabla f(x,y)=(\frac{x}{2},2y)$

Con X_0=(2,1):

1^{era} iteración:

$\nabla f(2,1)=(1,2)$ y entonces $x_1=(2,1)-d(1.2)=(2-d,1-2d)$.

Queda el problema: $Min\ \ f(x_1(d))=\frac{(2-d)^2}{4}+(1-2d)^2$, el que se resuelve derivando e igualando a cero: $f`(d)=\frac{17}{2}d-5=0$ y se obtiene d = 0.588.

Luego $x_1=(2,1)-0.588(1,2)=(1.412,-0.176)$

2^{da} iteración:

$\nabla f(1.412,-0.176)=(0.706,-0.352)$

$x_2=(1.412,-0.176)-d(0.706,-0.352)$

Queda el problema $Min\ \ f(x_2(d))=\frac{(1.412-0.706d)^2}{4}+(-0.176+0.352d)^2$

del cual se obtiene d = 0.863. Luego $x_2=(0.803,0.273)$

3^{ra} iteración:

$\nabla f(0.803,0.273)=(0.401,0.546)$

$x_3=(0.803,0.273)-d(0.401,0.546)$

Queda el problema $Min\ \ f(x_3(d))=\frac{(0.803-0.401d)^2}{4}+(0.273-0.546d)^2$

del cual se obtiene d = 1.026. Luego $x_3=(0.392,-0.287)$

4^{ta} iteración:

$\nabla f(0.392,-0.287)=(0.196,-0.574)$

$x_4=(0.392,-0.287)-d(0.196,-0.574)$

Queda el problema $Min\ f(x_4(d))=\frac{(0.392-0.196d)^2}{4}+(-0.287+0.574d)^2$ del que se obtiene d = 0.542. Luego $x_4=(0.286,0.024)$

5^{ta} iteración:

$$\nabla f(0.286, 0.024) = (0.143, 0.048)$$

$$x_5 = (0.286, 0.024) - d(0.143, 0.048)$$

Queda el problema $Min\ \ f(x_5(d)) = \dfrac{(0.286 - 0.143d)^2}{4} + (0.024 + 0.048d)^2$ del que se obtiene d = 1.447. Luego $x_5 = (0.079, -0.021) \approx (0,0)$

Es clarísimo que el mínimo global es (0,0), ya que $H(f) = \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix}$ es definido positivo.

Ahora tomemos como partida X_0 =(0.1):

1^{era} iteración:

$$\nabla f(0,1) = (0,2)$$

$$x_1 = (0,1) - d(0,2)$$

Queda el problema $Min\ \ f(x_1(d)) = (1-2d)^2$ del cual se obtiene d = ½

Luego $x_1 = (0,0)$.

Vemos que se llega al óptimo en una iteración, se debe a que la dirección de máximo decrecimiento pasaba por el óptimo (apuntaba al centro de las curvas de nivel elípticas), por lo que es importante elegir un buen punto de partida.

Problema 5:

Considere el siguiente problema de optimización $P)Min : f(x, y)$ y considere que usted inicia la búsqueda de una solución óptima en el punto X_0 y se mueve desde X_0 en la dirección de máximo descenso $-\nabla f(X_0)$. En todo punto $x(t), y(t)$ el gradiente de la función objetivo es igual a $\nabla f(x(t), y(t))$, y el vector de la dirección de movimiento es $-\nabla f(X_0)$. Demuestre que en el mínimo de $f(X)$, sobre esta dirección, estos dos vectores son ortogonales, es decir $\nabla f(x(t), y(t))$ con $-\nabla f(X_0)$.

Respuesta:

Para encontrar el mínimo se debe resolver $Min_t : f(x(t), y(t))$. Para encontrar un t^* óptimo, la condición de primer orden es:

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} = 0 .$$

Esto se puede rescribir con un producto punto:

$$\nabla f(x(t^*), y(t^*)) * \begin{bmatrix} \dfrac{dx(t^*)}{dt} \\ \dfrac{dy(t^*)}{dt} \end{bmatrix} = 0$$

Pero se tiene $X(t) = X_0 - t\nabla f(X_0)$. Luego:

$$\left[\frac{dx(t^*)}{dt}, \frac{dy(t^*)}{dt}\right] = -\nabla f(X_0)$$

Por lo que tenemos que los dos vectores son ortogonales en el punto mínimo.

Problema 6 (comente):

Verdadero o falso:

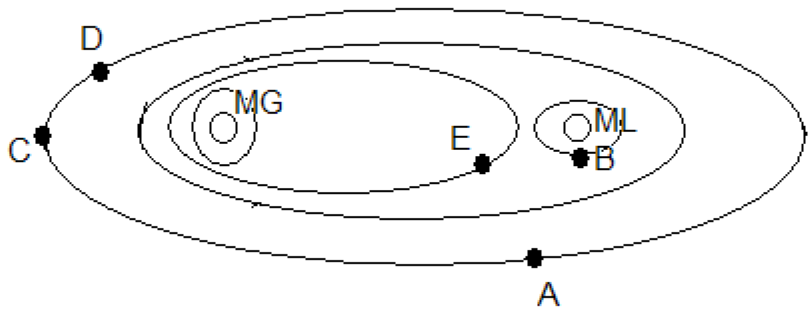
"En el método del gradiente la velocidad de convergencia depende exclusivamente de las formas de las curvas de la función objetivo."

Respuesta:

Falso, el método del gradiente al igual que el de newton dependen tanto de la forma de las curvas de nivel, como del punto de partida establecido X_0.

Problema 7 (comente):

Dado el dibujo de las curvas de nivel en dos dimensiones. ¿Que punto escogería para comenzar el método del gradiente?:



Respuesta:

Escogería C, ya que su dirección de máximo decrecimiento apunta directamente al punto óptimo. Convergería en una iteración.

Problema 8:

Considere el siguiente problema de minimización sin restricciones:

$$Min: \; x^2 + 2y^2 - 2xy - 2y$$

a) Justifique que el problema admite solución óptima
b) Realice 4 iteraciones del método del Gradiente a partir del punto (x,y) = (0,0)
c) Encuentre una expresión general para la iteración k.

Respuesta:

→Para la casa el desarrollo.

La expresión de la iteración "k" es:

$$x_k = \begin{pmatrix} \dfrac{2^{k/2}-1}{2^{\frac{k}{2}}} & \dfrac{2^{k/2}-1}{2^{\frac{k}{2}}} \end{pmatrix} \text{ si } k \text{ es par}$$

$$x_k = \begin{pmatrix} \dfrac{2^{(k-1)/2}-1}{2^{\frac{k-1}{2}}} & \dfrac{2^{(k+1)/2}-1}{2^{\frac{k+1}{2}}} \end{pmatrix} \text{ si } k \text{ es par}$$

Tomando límite $k \to \infty$ se obtiene como óptimo X = $(1 \quad 1)$.

c) MÉTODO DE NEWTON EN N-DIMENSIONES

→ Utiliza la misma idea que el método del gradiente ocupando el concepto de la dirección de descenso.
→ Parte de un punto X_0 y busca a través de la aproximación cuadrática (Taylor en n dimensiones) una dirección de descenso.

Paso Inicial:

- Se elige un vector inicial X_0

Paso Iterativo:

- Se construye $X_{k+1} = X_k - \overrightarrow{h(X_k)}$ con $h(X_k)$ dirección de descenso (o dirección de Newton) dada por:

$$\overrightarrow{h(X_k)} = H(f(X_k))^{-1} * \nabla f(X_k)^T$$

Término:

- Se prueba algún criterio de término.

Observaciones:

→ Este método es posible si la matriz Hessiana es invertible y es costoso invertirla. (Demora más en cada iteración del algoritmo)
→ Tiene convergencia cuadrática.
(Cada iteración avanza más rápido al óptimo y por lo tanto se necesitan menos)

Problema 9:

El método de newton en varias variables converge más rápido que el método del Gradiente. ¿Verdadero o falso?

Respuesta:

La afirmación es falsa. El método de newton es una aproximación de segundo orden, es decir tiene convergencia cuadrática, por lo que hace menos iteraciones que el método del gradiente. Sin embargo, en el método del gradiente las iteraciones, si bien son más, también son más rápidas, debido a que no se necesita calcular la matriz inversa del Hessiano en cada una. Luego la velocidad de convergencia es incierta y depende de cada problema específico.

Problema 10:

Explique el algoritmo base que hay detrás de los dos métodos numéricos visto en clase. ¿Qué diferencia fundamental hay entre ambos?

Respuesta:

Los dos métodos son el de newton y el del gradiente. Ambos métodos contemplan esta serie de pasos:

1.- Escoger un punto inicial.
2.- Determinar una dirección de movimiento.
3.- Moverse en esa dirección de acuerdo a algún criterio.
4.- Verificar el criterio de parada.

La diferencia fundamental entre ambos es la forma en que se determina la dirección de movimiento del paso 2 (Por la aproximación de Taylor de 2do orden o por la solución del problema de minimización unidimensional).

Problema 11:

Determine el mínimo de la siguiente función utilizando para ello el método de Newton y utilizando como punto inicial el (2,-2,1).

$$f(x,y,z) = x^2 + y^2 + 2z^2$$

Respuesta:

Primero se tiene que:

$$\nabla f(x,y,z) = [2x, 2y, 4z] \text{ y que: } H = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \rightarrow H^{-1} = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{bmatrix} \forall x,y,z$$

Iteración 1:

$$x_1 = x_0 - H^{-1}(f(x_o)) * \nabla f(x_0)$$

$$x_1 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{bmatrix} \begin{bmatrix} 4 \\ -4 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Iteración 2:

$$x_2 = x_1 - H^{-1}(f(x_1)) * \nabla f(x_1)$$

$$x_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Por lo que el punto $x = [0,0,0]$ es un punto crítico (analizando convexidad y existencia es óptimo)

OJO:

Se ha llegado al óptimo en una iteración debido a que el método converge de forma cuadrática y esta función es cuadrática.

OPTIMIZACIÓN NUMÉRICA CON RESTRICCIONES:

d) EL MÉTODO DE FRANK-WOLFE

Es un método numérico de búsqueda de solución, para un problema no lineal sujeto a restricciones del tipo:

$$\begin{aligned}&Min: f(\overrightarrow{X})\\&s.a:\\&g(\overrightarrow{X}) \leq b\\&x_i \geq 0,\ \ \forall i = 1,..,n\end{aligned}$$

Su gracia es que fracciona problemas no lineales (a veces muy difíciles) en problemas lineales (más fáciles) y problemas unidimensionales irrestrictos.

PASOS:

1- Determinar un punto $\overrightarrow{X_0}$.

Paso Iterativo:

2- Resolver el problema lineal (FASE I):

$$\begin{aligned}&Min: f(\overrightarrow{X_k}) + \nabla f(\overrightarrow{X_k})(\overrightarrow{X} - \overrightarrow{X_k})\\&s.a:\\&g(\overrightarrow{X}) \leq b\\&x_i \geq 0,\ \ \forall i = 1,..,n\end{aligned} \quad \Leftrightarrow \quad \begin{aligned}&Min: \nabla f(\overrightarrow{X_0})\overrightarrow{X}\\&s.a:\\&g(\overrightarrow{X}) \leq b\\&x_i \geq 0,\ \ \forall i = 1,..,n\end{aligned} \quad \Rightarrow \overrightarrow{Y_k} = \overrightarrow{X}*$$

3- Resolver el problema unidimensional irrestricto (FASE II):

$$\begin{aligned}&Min: f(\lambda\overrightarrow{X_k} + (1-\lambda)\overrightarrow{Y_k})\\&s.a:\\&0 \leq \lambda \leq 1\end{aligned} \quad \rightarrow \quad \lambda^* \quad \rightarrow \quad \overrightarrow{X_{k+1}} = \lambda^*\overrightarrow{X_k} + (1-\lambda^*)\overrightarrow{Y_k}$$

Test de Convergencia:

4- Evaluar la condición de convergencia.

$$f(\overrightarrow{X_k}) \approx f(\overrightarrow{X_{k+1}}) \quad \text{ó} \quad \left|\frac{f(X_{k+1}) - f(X_k)}{f(X_k)}\right| \leq \tau\%$$

OJO: ¡Asegura convergencia para problemas convexos!

Problema 12:

Resolver con el algoritmo de Frank – Wolfe hasta obtener un error menor al 5%:

$$Min: (x_1-1)^2+(x_2-7)^2+x_2$$
$$s.a.$$
$$x_1+2x_2\le 4$$
$$x_1,x_2\ge 0$$

Respuesta:

Paso inicial:

→ $X_0=\left[0,0\right]$ → f(X_0) = 50

→ $\nabla f(X)=[2(x_1-1),2(x_2-7)+1]$

Paso iterativo:

Iteración 1:

$h_0=\nabla f(X_0)*Y=\left[-2,-13\right]*\left[y_1,y_2\right]^T=-2y_1-13y_2$

Luego:

$$Min: -2y_1-13y_2$$
$$s.a:$$
$$y_1+2y_2\le 4$$
$$y_1,y_2\ge 0$$

La solución a este problema se puede obtener fácilmente. Por ejemplo con solución gráfica → $Y_0*=\left[0,2\right]$ → $h_0*=-26$ (f(Y_0) = 28)

Fabricamos X_1:

$X_1=(1-\lambda)X_0+\lambda Y_0=[0,2\lambda]$

Luego minimizamos el problema unidimensional:

$$Min: 1+(2\lambda-7)^2+2\lambda$$
$$0\le\lambda\le 1$$

Derivando e igualando a cero → $4(2\lambda-7)+2=0$ → $\lambda*=13/4$

Como λ es menor que 1 → $\lambda*=1$

$\boxed{\text{Luego: } X_1=\left[0,2\right]}$

$\boxed{f(X_1)=26}$

Iteración 2:

$h_1 = \nabla f(X_1) * Y = [-2,-9]*[y_1, y_2]^T = -2y_1 - 9y_2$

Luego:

$Min: -2y_1 - 9y_2$

$s.a:$

$y_1 + 2y_2 \leq 4$

$y_1, y_2 \geq 0$

$Y_1^* = [0,2]$ → $h_1^* = -26$ (f(Y_1) = 28)

Fabricamos X_2:

$X_2 = (1-\lambda)X_1 + \lambda Y_1 = [0,2]$

Luego no necesitamos minimizar el problema unidimensional y sabemos con certeza que el óptimo es X* = X_1 = $[0,2]$ (Se debe a que el problema es cuadrático).