

Programación NO Lineal (PNL)

Optimización sin restricciones

Ejemplos de los problemas que se aplica la programación NO Lineal:

- Problema de transporte con descuentos por cantidad : El precio unitario de transporte entre un origen y un destino es decreciente en función de la cantidad a transportar.
- Problema de flujo de cargas en un sistema eléctrico: Las pérdidas son no lineales .
- Problema de producción con elasticidad en el precio y/o en el coste:
Consideremos que los precios unitarios disminuyen cuando vendemos un número importante de productos, esto nos indica que la **renta marginal** disminuye cuando aumentan las ventas. Es el caso de reducir el precio de venta a partir de cierta cantidad vendida: cuando llevo vendido una cantidad importante le reduzco el precio de venta.

Extremos de la función : Los valores donde alcanza el máximo o el mínimo se llaman **extremos de la función**

Extremo global

La función objetivo $f(x_1; x_2; \dots; x_n)$ tiene un extremo global ; por ejemplo un máximo global, si es mayor o igual que cualquier otro valor de la función dentro del dominio de la misma.

Análogamente para mínimo si es menor o igual.

Extremo local

En este caso el extremo que alcanza la función es dentro de un entorno de un punto. Con el signo $\leq$; ó $\geq$ según corresponda a mínimo o máximo.

Podemos definir entonces que para Programación No Lineal un punto factible $\mathbf{X} = (x_1, x_2, \dots, x_n)$ es un máximo (mínimo) local si para un ε suficientemente pequeño,

cualquier punto factible $\mathbf{X}' = (x_1', x_2', \dots, x_n')$ tal que $|x_i - x_i'| < \varepsilon$ satisface

$$f(\mathbf{X}) \geq (\leq) f(\mathbf{X}')$$

Un punto que es un máximo o mínimo local se llama extremo local o relativo.

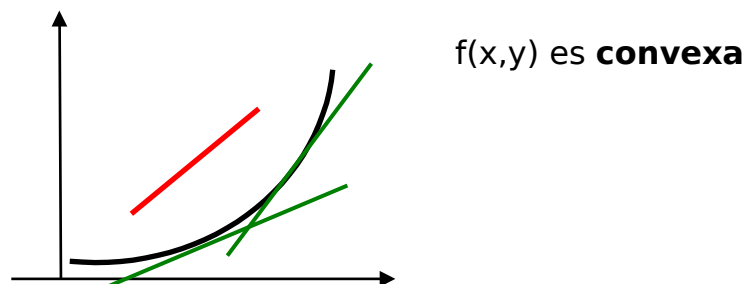
Funciones cóncavas y convexas

Función convexa:

Si hablamos de la función objetivo $f(x_1, x_2, \dots, x_n)$ podemos definir si se trata de una función Cóncava (cóncava hacia abajo) ó Convexa (cóncava hacia arriba)

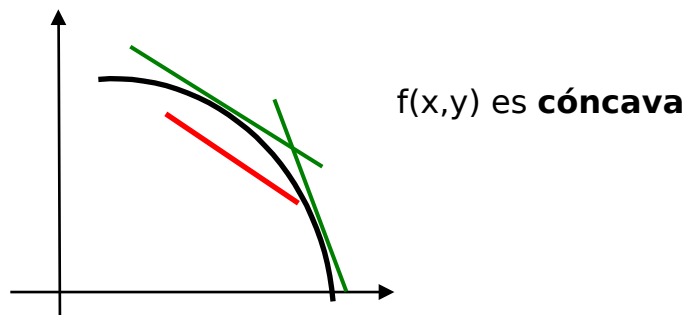
Para interpretar la definición se puede pensar en una función de 2 variables que podemos representar en el plano.

$f(x, y)$ es **convexa** si las cuerdas entre dos puntos de la función están por arriba de la curva que representa la función



Otra definición es que : las derivadas primeras son rectas que quedan por debajo de la función (rectas en verde)

$f(x, y)$ es **cóncava** si las cuerdas entre dos puntos de la función están por debajo de la curva que representa la función



En este caso las derivadas primeras son las rectas que permanecen por encima de la función (rectas en verde)

Esta condición la podemos anotar : f es **convexa** si

$$f[\lambda x_1 + (1 - \lambda) x_2] \leq \lambda f(x_1) + (1 - \lambda) f(x_2);$$

si el signo es $<$ es estrictamente convexa

donde λ es un número real entre 0 y 1 ; $0 \leq \lambda \leq 1$

análogamente f es **cóncava** si

$$f[\lambda x_1 + (1 - \lambda) x_2] \geq \lambda f(x_1) + (1 - \lambda) f(x_2)$$

si el signo es $>$ es estrictamente cóncava

Teorema 1

Se considera un PNL con región factible S convexa.

Entonces si el problema es de maximización (*minimización*) y la función f es estrictamente **cóncava** (*convexa*) en S , entonces cualquier máximo (*mínimo*) local del PNL es una solución óptima de este problema; es decir es un máximo global.

Teorema 2

Si $f''(X)$ existe para cualquier X en un conjunto convexo S , entonces $f(X)$ es una función **cóncava** (*convexa*) en S si y sólo si $f''(X) \leq 0$ (≥ 0) para todo X de S .

Teorema 3

Si $f(x_1, x_2, \dots, x_n)$ tiene derivadas parciales de segundo orden continuas para cada punto $X = (x_1, x_2, \dots, x_n)$ de un conjunto convexo S , entonces $f(X)$ es una función estrictamente **cóncava** (*convexa*) en S si y sólo si su matriz Hessiana es definida **negativa** (*positiva*).

Conclusiones:

El máximo (mínimo) global es difícil de encontrar.

Muchos métodos son locales.

Sólo bajo supuestos adicionales concavidad (convexidad) se puede garantizar que es global.

Dentro de un espacio solución (conjunto de inecuaciones) lineal convexo, **si $f(X)$ es cóncava**, el máximo obtenido es global es decir es el óptimo de $f(x)$, pues no existe otro mayor dentro del espacio solución.

Dentro de un espacio lineal convexo (conjunto de inecuaciones), **si $f(X)$ es convexa** el máximo obtenido es un máximo local. Es el mayor dentro de un subespacio del espacio solución pero No podemos asegurar que es el óptimo de $f(X)$.

Optimización sin restricciones en dimensión 1

Maximizar o minimizar $f(x)$ f continua en $[a,b]$

Para encontrar la solución óptima de este problema busquemos primero todos los máximos (o mínimos) locales.

La solución óptima será el máximo local (o mínimo) con el mayor (o menor) valor de $f(x)$.

Si $a=-\infty$ ó $b=\infty$ el problema puede no tener solución.

Hay tres tipos de puntos que pueden ser máximo o mínimos locales:

Puntos estacionarios de f ; es decir si $a < x < b$ y $f'(x)=0$

Puntos donde no existe $f'(x)$

Puntos extremos del intervalo $[a,b]$

Observaciones: definiciones de extremos relativos de Análisis Numérico permiten encontrar extremos en problemas sin restricción.

Definiciones

Primeras derivadas (gradiente)

Segundas derivadas (hessiano)

Optimización SIN restricciones

máximo de $f(x_1, x_2, \dots, x_n) = f(X)$

Si X^* es máximo global de $f(X)$ entonces $\frac{\partial f}{\partial x_j} = 0$ para $j= 1, 2, \dots, n$

es decir que la condición necesaria para que X^* sea un máximo de $f(X)$ cuando $f(X)$ es una función diferenciable es que las n derivadas parciales respecto de cada una de las variables sea igual a cero.

Si $f(X)$ es una función cóncava esta condición es también suficiente.

Condición necesaria y suficiente

Sea f cóncava y diferenciable en X^* ; X^* es un máximo si y sólo si

$\nabla f(X^*) = 0$; X^* es un máximo global si y sólo si f cóncava en toda la región.

Métodos numéricos para dimensión 1

Maximizar $f(x)$; función de una variable sin restricciones.

Donde $f(x)$ es diferenciable y cóncava.

La condición necesaria y suficiente para que una solución particular X^* sea óptima es que $\frac{\partial f}{\partial x} = 0$ en x^*

Una posibilidad es que :

1. En esta ecuación se pueda despejar x^* , entonces el problema queda resuelto.
2. Si es muy complicado despejar x^* pues la derivada es una función compleja entonces no es posible hallar la solución analítica y se recurre a un método numérico.

Método iterativo de búsqueda de la solución

Regla del punto medio (Plan de búsqueda de Balzano)

Se toma un $\varepsilon > 0$ que se utilizará para establecer el error aceptado.

Se encuentran $\underline{x}$; $\overline{x}$ iniciales por inspección ; encontrando valores en los cuales la derivada cambia de signo, de positiva a negativa, pues se trata de un máximo.

Se toma una solución de prueba inicial

$$x' = (\underline{x} + \overline{x}) / 2$$

Paso iterativo:

1. se calcula $\frac{df(x)}{dx}$ en x'
2. Si este valor calculado es ≥ 0 , se hace $\underline{x} = x'$
3. Si este valor calculado es ≤ 0 , se hace $\overline{x} = x'$
4. Se obtiene una nueva $x' = (\underline{x} + \overline{x}) / 2$
5. El proceso termina cuando $(\overline{x} - \underline{x}) \leq 2 * \varepsilon$
de esta manera la nueva x' se encuentra a una distancia de x^* menor que ε , que es el error aceptado.

Ejemplo : $f(x) = 6x - x^2$

$$f'(x) = 6 - 2x$$

Por inspección puedo adoptar $\underline{x} = 0$; $\overline{x} = 5$ dado que $f'(0) = 6$ y $f'(5) = -4$
Además :

$f''(x) = -2$; esto implica que $f(x)$ es cóncava en todos sus puntos

Adoptamos como tolerancia $\varepsilon = 0,1$

es decir que $(\overline{x} - \underline{x}) \leq 2 \cdot 0,1 = 0,2$

Iteración	$f'(x')$	$\underline{x}$	$\overline{x}$	x'	$(\underline{x} - \overline{x})$
0		0	5	$5/2 = 2,5$	5
1	$f'(5/2) > 0$	$5/2$	5	$15/4 = 3,5$	$5/2 = 2,5$
2	$f'(15/4) < 0$	$5/2$	$15/4$	$25/8 = 3,125$	$5/4 = 1,25$
3	$f'(25/8) < 0$	$5/2$	$25/8$	$45/16 = 2,825$	$5/8 = 0,62$
4	$f'(45/16) > 0$	$45/16$	$25/8$	$95/32 = 2,96$	0,31
5	$f'(95/32) > 0$	$95/32$	$25/8$	$195/64 = 3,04$	0,16

Luego $x' = 3,04$ es el máximo de $f(x)$

Optimización sin restricciones en dimension > 1

Maximizar o minimizar $f(x_1, x_2, \dots, x_n)$

con $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$

Suponemos que existen las derivadas parciales de primer y segundo orden de f y que son continuas.

Si X^* es un extremo local del PNL entonces para $i = 1, \dots, n$ se verifica

$$\frac{\partial f(X^*)}{\partial x_i} = 0$$

$$\partial x_i$$

es decir, X^* es un punto estacionario

Un punto estacionario que no es extremo local es un punto de silla.

Una posibilidad es que :

1. El conjunto de ecuaciones que resultan de establecer las derivadas parciales iguales a cero se puedan resolver analíticamente, entonces el problema queda resuelto.
2. Si es muy complicado o si no es posible hallar la solución analítica entonces se recurre a un método numérico.

El objetivo es alcanzar un punto en el que todas las derivadas parciales sean cero.

Es decir si extendemos el procedimiento de búsqueda de una dimensión es necesario emplear el valor de las derivadas parciales para seleccionar la dirección específica en la que conviene moverse. Esta selección implica el uso del gradiente de la FO.

Se supone que la FO es diferenciable, entonces tiene gradiente :

$\nabla f(X)$ en cada punto

En particular el gradiente en un punto específico X' es el vector cuyos elementos son las derivadas parciales respectivamente evaluadas en cada X' de manera que

$$\nabla f(X') = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right] \text{ en } X'$$

El significado del gradiente es que el cambio infinitesimal en X que maximiza la tasa a la que $f(X)$ aumenta, es el cambio que es proporcional a $\nabla f(X)$.

La dirección del gradiente se interpreta como la dirección del segmento de recta dirigido (flecha) que va

del origen $(0,0,\dots,0)$ al punto $(\partial f/\partial x_1, \partial f/\partial x_2, \dots, \partial f/\partial x_n)$

en donde $\partial f/\partial x_j$ se evalúa en cada x_j de X'

Entonces se puede decir que la tasa a la que aumenta $f(X)$ se maximiza, si los cambios (infinitesimales) en X se hacen en la dirección del gradiente . Como el objetivo es encontrar la solución factible que maximice $f(X)$ parece adecuado buscar en la dirección del gradiente.

Procedimiento de búsqueda eficiente debe ser moverse en la dirección del gradiente hasta que se alcance una solución óptima X^* , en la $\nabla f(X^*) = 0$

No resultaría práctico cambiar X continuamente en la dirección del gradiente , ya que esta serie de cambios requiere una reevaluación continua de las $\partial f/\partial x_j$ y del cambio de dirección en la trayectoria. Entonces la mejor forma de hacerlo es continuar el movimiento en una dirección fija a partir de un valor prueba inicial, sin detenerse hasta que $f(X)$ deje de aumentar. Este punto de detención sería el siguiente valor prueba por lo que se debe volver a calcular el gradiente para determinar la nueva dirección del movimiento.

Con este enfoque cada iteración incluye cambiar el valor prueba actual X'

se modifica $X' = X' + t^* \nabla f(X')$

donde t^* es el valor positivo que maximiza $f(X' + t^* \nabla f(X'))$

donde esta última expresión es la función f en donde

cada variable x_j del vector X'

$$x_j = x_j' + t \cdot \left(\frac{\partial f}{\partial x_j} \right) \quad \text{para } j = 1, 2, \dots, n$$

entonces x_j esta compuesta por valores constantes y t como única variable, y por lo tanto la función f se convierte en una expresión de una variable.

Las iteraciones de este proceso de búsqueda del gradiente continúan hasta que el gradiente sea igual a cero o dentro de la tolerancia ε admitida .

Ejemplo:

$$\text{maximizar } f(X) = 2.x.y + 2.y - x^2 - 2.y^2$$

Las derivada primera

$$\frac{\partial f}{\partial x} = 2.y - 2.x$$

$$\frac{\partial f}{\partial y} = 2.x + 2 - 4.y$$

Analizar la concavidad

$f(X)$ es cóncava:

$$\frac{\partial^2 f}{\partial x \partial x} = -2 < 0 \text{ Negativa}$$

$$\frac{\partial^2 f}{\partial y \partial y} = -4 < 0 \text{ Negativa}$$

$$X = (x, y) = (0, 0)$$

$$\frac{\partial f(0,0)}{\partial x} = 2.y - 2.x = 0$$

$$\frac{\partial f(0,0)}{\partial y} = 2.x + 2 - 4.y = 2$$

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

$$X' = X + t \nabla f(0, 0) = (0, 0) + t \cdot (0, 2) = (0, 2t)$$

$$f(0, 2t) = 2 \cdot 0 \cdot 2t + 2 \cdot 2t - 0 - 2 \cdot (2t)^2 = 4t - 8.t^2$$

$$t^* = \max f(t) = \max (4t - 8.t^2)$$

$$\frac{\partial f}{\partial t} = 0$$

$$\frac{\partial f}{\partial t}$$

$$= 4 - 16 t = 0$$

$$\frac{\partial f}{\partial t}$$

$$t^* = 1/4$$

$$X' = X + t \nabla f(0, 0) = (0, 0) + 1/4 \cdot (0, 2) = (0, 1/2)$$

$$\frac{\partial f(0,1/2)}{\partial x} = 2.y - 2.x = 1$$

$$\frac{\partial f(0,1/2)}{\partial y} = 2.x + 2 - 4.x = 0$$

$$\nabla f(X') = (1,0)$$

$$X' = X + t \nabla f(1,0) = (0,1/2) + t \cdot (1,0) = (t,1/2)$$

$$f(t,1/2) = 2.t.1/2 + 2.1/2 - t^2 - 2(1/2)^2 = -t^2 + t + 1/2$$

$$t^* = \max f(t) = \max (-t^2 + t + 1/2)$$

$$\frac{\partial f}{\partial t} = -2t + 1 = 0$$

$$t^* = 1/2$$

$$X' = X + t \nabla f(1,0) = (0,1/2) + 1/2 \cdot (1,0) = (1/2, 1/2)$$

$$\text{iteración } 0 \quad X' = (0, 0)$$

$$\text{iteración } 1 \quad X' = (0, 1/2)$$

$$\text{iteración } 2 \quad X' = (1/2, 1/2)$$

Ejemplo 2:

$$\text{maximizar } f(X) = -1/2.x^2 - 2.x.y - y^2 + y + 9$$

Las derivada primera

$$\frac{\partial f}{\partial x} = -x - 2y$$

$$\frac{\partial f}{\partial y} = -2.x - 2y + 1$$

Si resuelvo el sistema

$$\begin{aligned} -x - 2y &= 0 \\ -2.x - 2y + 1 &= 0 \end{aligned} \quad x=1 ; y=1/2$$

Analizar la concavidad

$f(X)$ es cóncava:

$$\frac{\partial^2 f}{\partial x \partial x} = -1 < 0 \text{ Negativa}$$

$$\frac{\partial^2 f}{\partial y \partial y} = -2 < 0 \text{ Negativa}$$

$$X = (x,y) = (0,0)$$

$$\frac{\partial f(0,0)}{\partial x} = -x - 2y = 0$$

$$\frac{\partial f(0,0)}{\partial y} = -2.x - 2y + 1 = 1$$

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

$$X' = X + t \nabla f(0, 0) = (0,0) + t \cdot (0,1) = (0,t)$$

$$f(0,t) = -1/2.0 - 2.0.t - (t)^2 + t + 9 = 9 + t - t^2$$

$$t^* = \max f(t) = \max (9 + t - t^2)$$

$$\frac{\partial f}{\partial t} = 0$$

$$\frac{\partial f}{\partial t}$$

$$\frac{\partial f}{\partial t} = 1 - 2t = 0$$

$$\frac{\partial f}{\partial t}$$

$$\mathbf{t^* = 1/2}$$

$$X' = X + t \nabla f(0, 0) = (0, 0) + 1/2 \cdot (0, 1) = (0, 1/2)$$

$$\frac{\partial f(0, 1/2)}{\partial x} = -x - 2y = -1$$

$$\frac{\partial f}{\partial x}$$

$$\frac{\partial f(0, 1/2)}{\partial y} = -2x - 2y + 1 = 0$$

$$\frac{\partial f}{\partial y}$$

$$\nabla f(X') = (-1, 0)$$

$$X' = X + t \nabla f(1, 0) = (0, 1/2) + t \cdot (-1, 0) = (-t, 1/2)$$

$$f(-t, 1/2) = -1/2 \cdot t^2 - 2 \cdot t \cdot 1/2 - (1/2)^2 + 1/2 + 9 = -1/2 t^2 - t + 37/4$$

$$t^* = \max f(t) = \max (-1/2 t^2 - t + 37/4)$$

$$\frac{\partial f}{\partial t} = -t - 1 = 0$$

$$\frac{\partial f}{\partial t}$$

$$\mathbf{t^* = -1}$$

$$X' = X + t \nabla f(1, 0) = (0, 1/2) + -1 \cdot (-1, 0) = (1, 1/2)$$

$$\text{iteración } 0 \quad X' = (0, 0)$$

$$\text{iteración } 1 \quad X' = (0, 1/2)$$

$$\text{iteración } 2 \quad X' = (1, 1/2)$$