

IN3701 – Modelamiento y Optimización
Complemento Auxiliar 3**Programación separable:**

La programación separable es un método de programación no lineal. Es un método importante, porque permite aproximar un modelo no lineal a un modelo de programación lineal, con un nivel de ajuste arbitrario.

Sea un modelo de programación no lineal como sigue:

$$\text{minimizar}\{f(x)\}$$

con dos puntos importantes: La función f y las restricciones son *separables* aditivamente y la variable X_j está acotada inferiormente por cero y superiormente por u_j ($j=\{1,\dots,n\}$).

La función separable $f(x)$ es separable si se puede expresar como la suma de funciones donde cada sumando contiene una sola variable X_j :

$$f(x) = \sum_{j=1}^n f_j(x_j)$$

Entonces, el problema no lineal nos queda de la siguiente forma:

$$\begin{aligned} &\text{Minimize} \quad \sum_{j=1}^n f_j(x_j) \\ &\text{s.a.} \\ &0 \leq x_j \leq u_j, \quad j=1,\dots,n \end{aligned}$$

La gracia de esta formulación es que las “no linealidades” son independientes. Esto junto con el hecho de que la variable de decisión X_j es acotada, nos permite encontrar una aproximación lineal a este problema.

(en la auxiliar partimos desde aquí, pero deben considerar los supuestos anteriores)

Linealización

Consideremos la función no lineal general $f(x)$ como se muestra en la figura 1. Para formar una aproximación lineal por tramos, debemos primero dividir el eje x en $r+1$ puntos distintos, no necesariamente equidistantes, asumiendo que queremos formar r tramos lineales. Sean estos puntos $\bar{x}_0, \dots, \bar{x}_r$ y sea $f_k = f(\bar{x}_k)$ para $k=\{0,\dots,r\}$

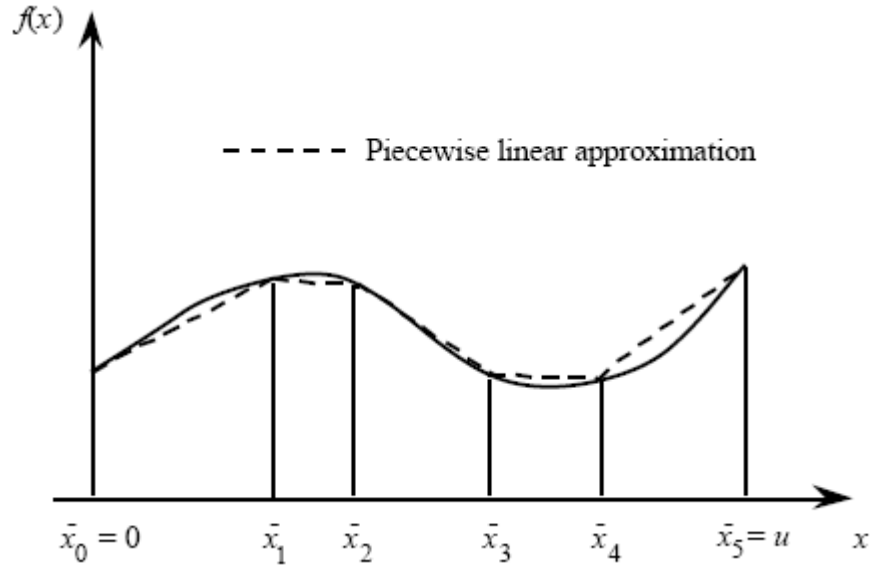


Figura 1: función no lineal f y su aproximación lineal por tramos

Notar que cualquier valor de x que esté entre dos de los puntos $\bar{x}_k$ definidos anteriormente, podemos decir que:

$$x = \alpha \bar{x}_{k+1} + (1 - \alpha) \bar{x}_k \text{ or } x - \bar{x}_k = \alpha (\bar{x}_{k+1} - \bar{x}_k) \text{ for } 0 \leq \alpha \leq 1. \quad (*)$$

Esto es fácil de comprobar, basta con probar valores extremos de α . Si $\alpha=0$ entonces $x = \bar{x}_k$, pero si $\alpha=1$ entonces $x = \bar{x}_{k+1}$. Cualquier α con valores intermedios genera cualquier punto en la recta entre $\bar{x}_k$ y $\bar{x}_{k+1}$. Notar que los $\bar{x}_k$ son datos, ustedes los eligieron "al ojo", mientras que α es una variable de decisión (pues x depende de este α).

Ahora, podemos representar el tramo de la función lineal que está entre $\bar{x}_k$ y $\bar{x}_{k+1}$, escribiendo la ecuación de la recta:

$$\hat{f}(x) = f_k + \frac{f_{k+1} - f_k}{\bar{x}_{k+1} - \bar{x}_k} (x - \bar{x}_k) = \alpha f_{k+1} - (1 - \alpha) f_k \text{ for } 0 \leq \alpha \leq 1$$

Notar que usó la ecuación (*) para llegar a la última expresión desde la ecuación de la recta.

Renombremos ahora el parámetro α . Sea $\alpha = \alpha_{k+1}$ y $(1 - \alpha) = \alpha_k$. Así, para $\bar{x}_k \leq x \leq \bar{x}_{k+1}$ se tiene que

$$x = \alpha_{k+1} \bar{x}_{k+1} + \alpha_k \bar{x}_k \text{ and } \hat{f}(x) = \alpha_{k+1} f_{k+1} + \alpha_k f_k$$

donde $\alpha_k + \alpha_{k+1} = 1$ y ambos son positivos.

Generalizando este procedimiento para cubrir toda la función (y no solo el tramo donde $\bar{x}_k \leq x \leq \bar{x}_{k+1}$), se obtiene la siguiente expresión:

$$x = \sum_{k=0}^r \alpha_k \bar{x}_k, \quad \hat{f}(x) = \sum_{k=0}^r \alpha_k f_k, \quad \sum_{k=0}^r \alpha_k = 1, \quad \alpha_k \geq 0, \quad k=0, \dots, r$$

notando que de al menos un α_k y no más de dos pueden ser mayores que cero. Además, necesitamos que ambos α_k que sean distintos de cero DEBEN ser adyacentes (i.e. su índice debe tener exactamente un espacio de diferencia). Por ejemplo, si α_5 es distinto de cero, solo α_4 o α_6 pueden ser distintos de cero, todo el resto debe ser igual a cero. Deben ser adyacentes porque así estamos justo tomando un tramo de la curva (miren la figura y debería quedarles claro). Si no fueran adyacentes estaríamos tomando un tramo que no pertenece a la aproximación lineal que dibujamos. En otras palabras, si el criterio de adyacencia no se satisface, la aproximación lineal de $f(x)$ NO ES la que queríamos, la que dibujamos.

Por lo tanto, nos falta definir un conjunto de restricciones que nos aseguren que los alfa sean adyacentes. Esto se define de la siguiente forma:

$$\begin{aligned} \alpha_1 &\leq y_1 \\ \alpha_k &\leq y_{k-1} + y_k \quad \forall k = \{2, \dots, r-1\} \\ \alpha_r &\leq y_{r-1} \\ \sum_k y_k &= 1 \end{aligned}$$

Donde Y_k es una variable binaria que vale 1 si $\bar{x}_k \leq x \leq \bar{x}_{k+1}$ y cero sino. Aquí fue donde nos enredamos en la auxiliar. Pero así está bien, esos son los índices correctos. La gracia de estas restricciones es que, como la suma de la variable Y_k vale 1, entonces solo una variable Y vale 1 y todas las otras valen cero. si prueban poner números en vez del índice K y asumen que Y_k vale 1 para un k fijo, verán que se cumple lo que queremos. Si por ejemplo $Y_3=1$, $\alpha_k \leq 1$ sólo para $k=3$ y $k=4$ y $\alpha_k=0$ para el resto. Como además la suma de los alfa es igual a 1 entonces estamos ok, tenemos dos alfas adyacentes distintos de cero (o al menos uno de ellos lo es).

Y con eso estamos listos, el problema original

$$\begin{aligned} &\underset{n}{\text{Minimize}} \quad \sum_{j=1}^n f_j(x_j) \\ &\text{s.a.} \\ &0 \leq x_j \leq u_j, \quad j=1, \dots, n \end{aligned}$$

Nos quedaría:

$$\begin{aligned}
& \min \hat{f}(x) \\
& \text{s.a.} \\
& x = \sum_{k=0}^r \alpha_k x_k \\
& \hat{f}(x) = \sum_{k=0}^r \alpha_k f_k \\
& \sum_{k=0}^r \alpha_k = 1 \\
& \alpha_1 \leq y_1 \\
& \alpha_k \leq y_{k-1} + y_k \quad \forall k = \{2, \dots, r-1\} \\
& \alpha_r \leq y_{r-1} \\
& \sum_k y_k = 1
\end{aligned}$$

Y eso es todo.

Ahora... ¡OJO! Si la función a aproximar por tramos es convexa y el problema es de minimización, entonces SIEMPRE los alfa que elija el modelo van a ser adyacentes y, por lo tanto, no hace falta definir las restricciones asociadas. En otras palabras, para un problema de minimización y función convexa, nos basta con decir que

$$\begin{aligned}
& \min \hat{f}(x) \\
& \text{s.a.} \\
& x = \sum_{k=0}^r \alpha_k x_k \\
& \hat{f}(x) = \sum_{k=0}^r \alpha_k f_k \\
& \sum_{k=0}^r \alpha_k = 1
\end{aligned}$$

Esta última es la que se usa en el problema 3 de la auxiliar, que van a ver que es bien sencillo. Lo importante es que entiendan de donde viene todo esto y no quedarse solo con el ejercicio numérico, ya que reemplazar valores es la parte fácil.

Asimismo, si la función es cóncava pero el problema es de maximización, ocurre lo mismo, los alfa serán siempre adyacentes y no hace falta poner ese set de restricciones. Queda propuesto (a conciencia) ver por qué esto es así. No les pido que lo demuestren, con que lo vean dibujándolo basta ;)

Dudas y/o comentarios a:
André Carboni
andre@carboni.cl