AMPLIACIÓN DE ESTADÍSTICA Y OPTIMIZACIÓN

Francisco J. Aragón Artacho
francisco.aragon@ua.es

Dpto. Matemáticas, Universidad de Alicante
Curso 2022-2023

Máster en Ciencia de Datos
Universidad de Alicante

Formulación continua de problemas con variables binarias

Ejemplo (Inversión de capital)

Supongamos que queremos invertir 500€ en los siguientes proyectos:

Proyecto	1	2	3	4	5	6	7
Coste	100	20	150	50	50	150	150
Beneficio	300	50	350	110	100	250	200

¿Cómo elegir la selección óptima?

Una variable 0-1 (también llamada binaria o Booleana) se puede convertir en continua mediante la equivalencia

$$x \in \{0, 1\} \Leftrightarrow x\,(1 - x) = 0.$$

(Análogamente, tenemos que $x \in \{-1, 1\} \Leftrightarrow x^2 = 1$.)

Formulación continua de problemas con variables enteras

Si $x \in [0, M]$ y se requiere $x \in \mathbb{Z}$ (entero), podemos escribir x en binario:

$$x = y_0 + y_1 2 + \ldots + y_p 2^p,\ y_i \in \{0, 1\},\ i = 0, 1, 2, \ldots, p,$$

si tomamos p suficientemente grande para asegurar que $\underbrace{1 + 2 + 2^2 + \cdots + 2^p}_{x} \geq M$.

$$x = 1 + \cancel{2} + \cancel{2^2} + \cancel{\ldots} + \cancel{2^p}$$

$$2x = \cancel{2} + \cancel{2^2} + \cancel{2^3} + \ldots \cancel{+ 2^p} + 2^{p+1}$$

$$x = 2^{p+1} - 1 \rightarrow 2^{p+1} = x + 1$$

$$(p+1)\log(2) = \log(2^{p+1}) = \log(x+1)$$

$$p + 1 = \frac{\log(x+1)}{\log 2} \rightarrow \boxed{p = \frac{\log(x+1)}{\log 2} - 1}$$

Si $M = 3244$)

$$p > \frac{\log(3245)}{\log(2)} - 1$$

$$= 10.66$$

$$\rightarrow \boxed{p > 11}$$

Formulación continua de problemas con variables enteras

Si $x \in [0, M]$ y se requiere $x \in \mathbb{Z}$ (entero), podemos escribir x en binario:

$$x = y_0 + y_1 2 + \ldots + y_p 2^p, \; y_i \in \{0, 1\}, \; i = 0, 1, 2, \ldots, p,$$

si tomamos p suficientemente grande para asegurar que $1 + 2 + 2^2 + \cdots + 2^p > M$.
Si $x \in [-M, M]$ y se requiere $x \in \mathbb{Z}$, podemos escribir

$$x = y - z, \; 0 \leq y \leq M, \, 0 \leq z \leq M, \; y, z \in \mathbb{Z}.$$

Sustituyendo x por las variables $y_0, \ldots, y_p$ y $z_0, \ldots, z_p$ mediante

$$x = (y_0 - z_0) + (y_1 - z_1)\, 2 + \ldots + (y_p - z_p)\, 2^p,$$

añadiendo las restricciones

$$\left\{ \begin{array}{l} y_i(1 - y_i) = 0, \; i = 0, \ldots, p \\ z_i(1 - z_i) = 0, \; i = 0, \ldots, p \end{array} \right\}.$$

Reformulación de restricciones

Podemos transformar un problema de optimización con igualdades en uno que solo contenga desigualdades, pues

$$h(x) = 0 \Leftrightarrow [h(x) \geq 0 \text{ y } h(x) \leq 0] \, .$$

También podemos convertir una desigualdad en una igualdad, añadiendo una nueva variable:

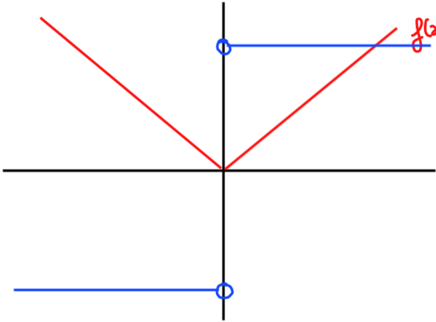
$$g(x) \leq 0 \Leftrightarrow g(x) + v^2 = 0,$$

$$\begin{cases} g(x) + u = 0 \\ u \geq 0 \end{cases}$$

donde v toma valores en $\mathbb{R}$.

Eliminación de valores absolutos en la función objetivo

La diferenciabilidad es una propiedad deseable de la función objetivo. Cuando aparece un valor absoluto en la función objetivo, ésta normalmente no será diferenciable.



$$f(x) = |x| = \begin{cases} x, & \text{si } x \geq 0 \\ -x, & \text{si } x < 0 \end{cases}$$

$$f'(x) = \begin{cases} 1, & \text{si } x > 0 \\ -1, & \text{si } x < 0 \end{cases}$$

Eliminación de valores absolutos en la función objetivo

La diferenciabilidad es una propiedad deseable de la función objetivo. Cuando aparece un valor absoluto en la función objetivo, ésta normalmente no será diferenciable.

Para simplificar, supongamos que P es un problema irrestringido con solo dos variables, x e y, y que la función objetivo f tiene la forma $f(x,y) = g(x,y) + |x|$:

$$P: \quad \text{Min} \quad g(x,y) + |x|.$$

$x = u - v,\ u \geq 0,\ v \geq 0$

$x = 5,\ v = 0,\ u = 5 \to u - v = 5 = x$

$v = 1,\ u = 6 \to u - v = 5 = x$

$x = -3 \to u = 0,\ v = 3$

$u - v = -3 = x$

$$|x| = \begin{cases} x, & \text{si } x \geq 0 \\ -x, & \text{si } x < 0 \end{cases} = \begin{cases} u - v, & \text{si } u \geq v \\ v - u, & \text{si } v > u \end{cases} \leq u + v$$

P_1: Min $g(u-v, y) + u + v =: h(u, v, y)$

s.a $u \geq 0,\ v \geq 0$

$\boxed{\text{Si } (u,v,y) \text{ es sol. óptima de } P_1 \to u = 0 \vee v = 0}$

D: RA: No se cumple $u = 0 \vee v = 0 \to u > 0,\ v > 0$. Sea $\delta := \min\{u, v\} > 0$

siendo (u,v,y) sol. óptima de P_1

Vamos a ver que $(u-\delta, v-\delta, y)$ es mejor que (u,v,y)

$u - \delta \geq u - u = 0,\ v - \delta \geq 0$

$g(u - \cancel{\delta} - (v - \cancel{\delta}), y) + (u-\delta) + (v-\delta) = g(u-v, y) + u + v - \underbrace{2\delta}_{> 0} < g(u-v, y) + u + v \to$ Como $(u-\delta, v-\delta, y)$ es factible, se tiene que (u,v,y) no puede ser sol. óptima $\lightning$ $\square$

Eliminación de valores absolutos en la función objetivo

La diferenciabilidad es una propiedad deseable de la función objetivo. Cuando aparece un valor absoluto en la función objetivo, ésta normalmente no será diferenciable.

Para simplificar, supongamos que P es un problema irrestringido con solo dos variables, x e y, y que la función objetivo f tiene la forma $f(x, y) = g(x, y) + |x|$:

$$P: \quad \text{Min} \quad g(x, y) + |x|.$$

Considerar el problema

$$\begin{array}{lll} P_1: & \text{Min} & g(u - v, y) + u + v \\ & \text{s.t.} & u, v \geq 0. \end{array}$$

Eliminación de valores absolutos en la función objetivo

La diferenciabilidad es una propiedad deseable de la función objetivo. Cuando aparece un valor absoluto en la función objetivo, ésta normalmente no será diferenciable.

Para simplificar, supongamos que P es un problema irrestringido con solo dos variables, x e y, y que la función objetivo f tiene la forma $f(x,y) = g(x,y) + |x|$:

$$P: \quad \text{Min} \quad g(x,y) + |x|.$$

Considerar el problema

$$\begin{array}{lll} P_1: & \text{Min} & g(u-v,y) + u + v \\ & \text{s.t.} & u, v \geq 0. \end{array}$$

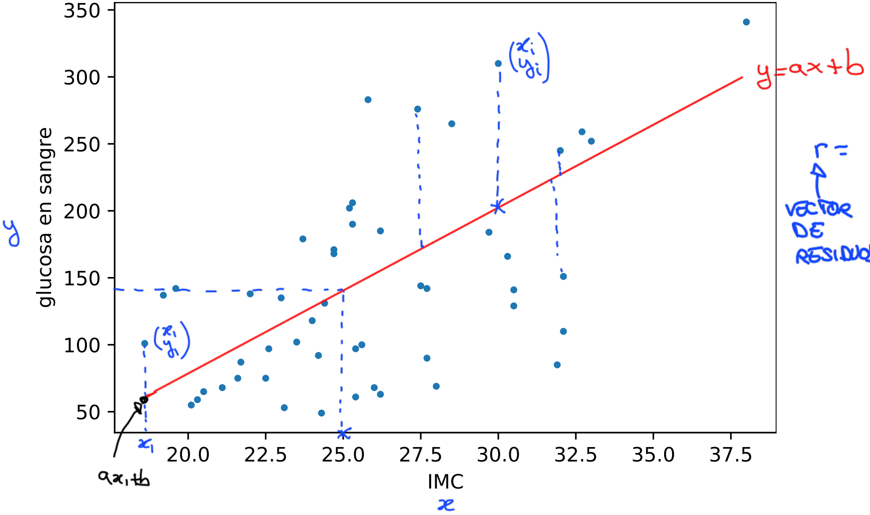
Por tanto, las soluciones óptimas de P y P_1 son las mismas, en el sentido de que si $\overline{x}$ es óptimo para P, definiendo

$$\overline{u} := \begin{cases} \overline{x}, & \overline{x} \geq 0 \\ 0, & \overline{x} < 0 \end{cases} \quad \text{and} \quad \overline{v} = \begin{cases} 0, & \overline{x} \geq 0 \\ -\overline{x}, & \overline{x} < 0, \end{cases}$$

tenemos que $(\overline{u}, \overline{v})$ es óptimo para P_1, y si $(\tilde{u}, \tilde{v})$ es óptimo para P_1, entonces $\tilde{x} = \tilde{u} - \tilde{v}$ es óptimo para P.

Regresión lineal

m individuos



$$r = \begin{pmatrix} ax_1+b-y_1 \\ ax_2+b-y_2 \\ \vdots \\ ax_i+b-y_i \\ \vdots \\ ax_m+b-y_m \end{pmatrix} \in \mathbb{R}^m$$

VECTOR DE RESIDUOS

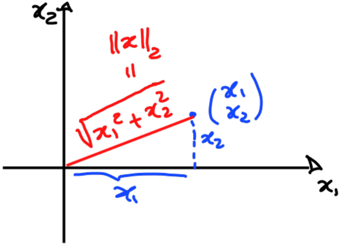
Regresión lineal

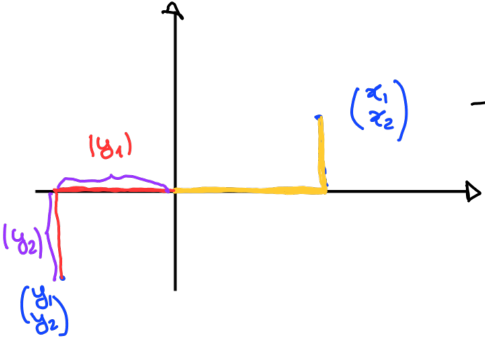
La normas $\|x\|_2$ (o Euclídea), $\|x\|_1$ (o de Manhattan) y $\|x\|_\infty$ (o de Chebyshev):

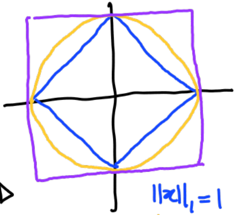
$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{R}^m$$

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_m^2} \qquad \|x\|_\infty = \text{máx}\{|x_1|, |x_2|, \dots, |x_m|\}$$

$$\|x\|_1 = |x_1| + |x_2| + \dots + |x_m| \qquad \|x\|_p = \left(|x_1|^p + |x_2|^p + \dots + |x_m|^p\right)^{1/p}$$

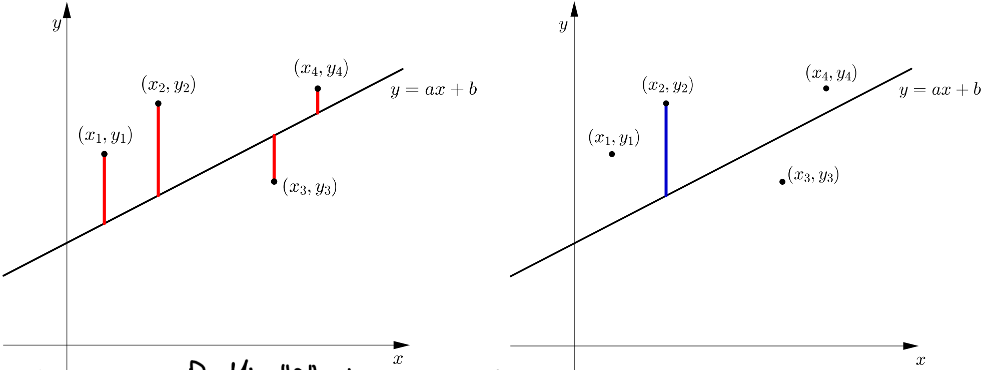






Regresión lineal

La normas $\|x\|_2$ (o Euclídea), $\|x\|_1$ (o de Manhattan) y $\|x\|_\infty$ (o de Chebyshev):



$$r = \begin{pmatrix} ax_1+b-y_1 \\ ax_2+b-y_2 \\ \vdots \\ ax_m+b-y_m \end{pmatrix} \in \mathbb{R}^m$$

$$P_1 : \min_{a,b\in\mathbb{R}} \|r\|_1 = |ax_1+b-y_1| + \dots + |ax_m+b-y_m|$$

$$P_2 : \min_{a,b\in\mathbb{R}} \|r\|_2^2 = (ax_1+b-y_1)^2 + \dots + (ax_m+b-y_m)^2$$

$$P_3 : \min_{a,b\in\mathbb{R}} \|r\|_\infty = \max\{|ax_1+b-y_1|, \dots, |ax_m+b-y_m|\}$$

Regresión lineal

$$P_1: \underset{a,b\in\mathbb{R}}{\text{Min}}\ |ax_1+b-y_1| + |ax_2+b-y_2| + \dots + |ax_m+b-y_m|$$ ← 2 variables

$u=\begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix}$ $v=\begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$

$$\tilde{P}_1: \text{Min}\ u_1+v_1+u_2+v_2+\dots+u_m+v_m$$

s.a
$$ax_1+b-y_1=u_1-v_1$$
$$ax_2+b-y_2=u_2-v_2$$
$$\vdots$$
$$ax_m+b-y_m=u_m-v_m$$
$$u_1\geq 0, u_2\geq 0, \dots, u_m\geq 0$$
$$v_1\geq 0, v_2\geq 0, \dots, v_m\geq 0$$

2m+2 variables

Es un problema de programación lineal (PL) porque la función objetivo y las restricciones son lineales

$$(1_m^T, 1_m^T, 0, 0)\begin{pmatrix} u \\ v \\ a \\ b \end{pmatrix} = 1_m^T u + 1_m^T v = u_1+\dots+u_m+v_1+\dots+v_m$$

$$a\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} + \begin{pmatrix} b \\ \vdots \\ b \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} - \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$$

$$ax+b1_m-y=u-v$$

$$\tilde{P}_1: \text{Min}\ (1_m^T, 1_m^T, 0, 0)\begin{pmatrix} u \\ v \\ a \\ b \end{pmatrix}$$

s.a
$$ax+b1_m-u+v=y$$
$$u\geq 0, v\geq 0$$