Guía Módulo 1. Introducción a la Investigación Operativa para el Data Science

Presentación

Las técnicas de la Investigación Operativa, y en particular la Programación Lineal, son a menudo utilizadas con el objetivo de encontrar soluciones óptimas a problemas que aparecen en ámbitos empresariales diferentes como la logística, el marketing o los recursos humanos.

En este primer módulo nos introduciremos en los aspectos básicos y las aplicaciones de la Programación Lineal, y de manera simultánea presentaremos algunos aspectos generales de la IO que nos ofrecerán una mejor comprensión del enorme potencial de aplicación de estas técnicas en la sociedad actual.

Objetivos específicos

Como objetivos específicos para este módulo destacamos los siguientes:

- Entender los conceptos básicos de un problema de programación lineal (PL): función lineal, variables de decisión, objetivo, restricciones, región factible, vértice, solución factible y solución factible óptima.
- Conocer unos primeros ejemplos de modelos de programación lineal (producción, transporte, inventario, asignación de personal...) que permiten plantear los problemas de PL.
- Aprender a comprobar la coherencia interna de un modelo de PL.
- Saber resolver gráficamente un problema de dos variables mediante los métodos de la línea de isobeneficio (isocoste) y de los vértices; esta resolución gráfica cuenta con la ayuda del software Excel.
- Clasificar correctamente los PLs según la tipología de sus soluciones.

Conocimientos previos

El estudio de los contenidos asociados a este módulo no requiere otros conocimientos previos específicos que no sean los conocimientos básicos de un primer curso de matemáticas universitarias.

Conceptos fundamentales

Este módulo, llamado "Introducción a la Investigación Operativa para el *Data Science*", empieza con un apartado totalmente introductorio que nos explica las diferentes técnicas de la Investigación Operativa (IO) y nos presenta los diferentes problemas que pueden tratarse mediante métodos de la IO. En consecuencia, se recomienda leer este apartado de forma esmerada prestando atención a las ideas fundamentales que figuran en él. No hay que perder de vista, sin embargo, que lo que pretendemos en nuestra asignatura es introducirnos en las técnicas de optimización lineal. Se recomienda la lectura adicional de los capítulos iniciales de cualquiera de los dos libros de la bibliografía básica (Hillier o Taha).

En el apartado 2 se exponen una serie de ejemplos de formulación de problemas de optimización lineal. Los complementos de este apartado se encuentran en la miscelánea de ejercicios (colección de problemas resueltos) del módulo 1, que forman parte de los recursos complementarios del curso.

El siguiente apartado de este módulo se llama "Modelos de Programación Lineal". Se presentan en él los **conceptos básicos de la optimización lineal**.

El apartado 4, llamado "Resolución gráfica de problemas lineales", es eminentemente práctico y nos tiene que permitir consolidar los conceptos básicos expuestos en el apartado 3. Se desarrolla en él **la resolución gráfica de modelos lineales** con dos variables. El hecho de que un problema tenga sólo dos variables permite su representación gráfica en un plano de ejes coordenados. En el estudio de este apartado es recomendable ir traduciendo todos los conceptos básicos de la optimización lineal (función objetivo, variables, restricciones, conjunto de soluciones posibles, tipo de soluciones...) a conceptos geométricos, que permiten un tratamiento más intuitivo. La importancia de este apartado radica en que una buena comprensión de los conceptos básicos en problemas con dos variables nos permitirá generalizar más adelante estos conceptos a problemas con muchas más variables, cuya representación geométrica es imposible.

En el manual de Excel módulo 1 encontraremos una descripción detallada de cómo utilizar este programa para la resolución gráfica de un PL. Se puede utilizar cualquier otro programa para resolver este tipo de problemas o incluso resolverlo "a mano".

Es muy recomendable la consulta de los apuntes de la asignatura, los recursos complementarios (miscelánea y manual de Excel) y de los libros de la bibliografía básica de la asignatura.

Ejercicios tipo

En este documento os presentamos cinco ejercicios tipo resueltos que no pretenden ser una lista exhaustiva de los tipos de problemas que pueden plantearse relacionados con los contenidos del módulo 1. Los cuatro primeros son ejercicios tipo de modelización de problemas mientras que el último es un ejercicio tipo resuelto utilizando la hoja de cálculo Excel. Por lo tanto, se hace necesario y recomendable consultar más ejemplos y ejercicios resueltos en los apuntes de la asignatura, en los recursos complementarios y en los libros de la bibliografía básica.

EJERCICIO 1

Una empresa de logística tiene un almacén donde puede guardar tres tipos de productos diferentes: A, B y C. La demanda por parte de sus clientes de cada uno de ellos es de, como mínimo, 70, 50 y 100 unidades respectivamente. Además, se conoce que la demanda estimada total para los tres tipos de producto es exactamente de 300 unidades. El coste variable unitario de mantenimiento de existencias (o almacén) para cada tipo de producto es de 100 €, 150 € y 200 €, respectivamente. También hay unos costes fijos debidos al servicio de vigilancia del almacén. Estos costes fijos son, en total, de 2.000 €. A partir de algunos estudios realizados por la empresa se pretende que el coste variable total de los productos A y B sea, como máximo, de un 50% del coste variable total de los tres productos. También se pretende que el coste variable del producto A sea, como mínimo, de un 30% del coste variable total de los tres productos. Por último, los precios de venta de cada uno de los productos son de 1.000 €, 1.250 € y 1.500 €, respectivamente. La empresa desea decidir cuántas unidades de cada producto debe tener en el almacén con el fin de maximizar el beneficio de su actividad comercial.

Resolución:

- Definición de variables:
  X_j: número de unidades almacenadas de cada uno de los productos A, B y C
- Restricciones de demanda: $X_a \geq 70$; $X_b \geq 50$; $X_c \geq 100$; $X_a + X_b + X_c = 300$
- Restricciones de costes:
  Dado que $100X_a + 150X_b + 200X_c$ es el coste variable total, entonces
  - $100X_a + 150X_b \leq 0.5(100X_a + 150X_b + 200X_c)$, de donde $50X_a + 75X_b - 100X_c \leq 0$
  - $100X_a \geq 0.3(100X_a + 150X_b + 200X_c)$, de donde $70X_a - 45X_b - 60X_c \geq 0$
- Función objetivo: restamos ingresos menos costes

Max $z = 900Xa + 1100Xb + 1300X_c - 2000$. Vemos que hay un término constante (–2.000) que hay que extraer fuera de la función a optimizar. Entonces la función lineal que optimizamos es Max $z^* = 900X_a + 1100X_b + 1300X_c$. Una vez optimizada esta función z* le restaremos2.000 (los costes fijos) al valor óptimo y obtendremos el valor óptimo de la función z.

- Por lo tanto, el modelo del PL es:
  Max $z = 900X_a + 1100X_b + 1300X_c$
  $X_a \geq 70$
  $X_b \geq 50$
  $X_c \geq 100$
  $X_a + X_b + X_c = 300$
  $50X_a + 75X_b - 100X_c \leq 0$
  $70X_a - 45X_b - 60X_c \geq 0$
  $X_i \in Z^+$ (las variables toman valores en el conjunto de los números enteros no negativos)

EJERCICIO 2

Una empresa produce coches y camiones. La empresa dispone de 4 departamentos (estampado, planchista, acabados coche, acabados camión) con restricciones de disponibilidad de horas de mano de obra. En la tabla anexa se muestra cuántas horas de mano de obra necesita cada coche o camión en cada departamento y la disponibilidad total de horas.

	Depto. I	Depto. II	Depto. III	Depto. IV
Coche	7	1	1	0
Camión	5	2	0	1
Total horas mano de obra disponibles en departamento	175.000	33.000	22.500	15.000

El precio de venta de los coches es de 300.000 por unidad y el de los camiones de 250.000. Los costes variables de los coches son de 50.000 por unidad y los de los camiones, 40.000. Los costes fijos de la empresa son de 50.000. Se trata de formular un PL con el fin de maximizar los beneficios.

Resolución:

- Definición de variables: X_j: número de coches y camiones producidos. Toman valores no negativos y enteros.

- Restricciones para cada departamento (horas disponibles):
  Depto. I: $7x_1 + 5x_2 \leq 175.000$
  Depto. II: $x_1 + 2x_2 \leq 33.000$
  Depto. III: $x_1 \leq 22.500$
  Depto. IV: $x_2 \leq 15.000$

- Función objetivo:
  Max $z=250.000*x_1+210.000*x_2-50.000$; dado que incluye un término constante (¡los costes fijos!) y no es lineal, la función que optimizaremos será Max $z^*=250.000*x_1+210.000*x_2$. Una vez encontrado el valor óptimo de z* le restaremos el término constante 50.000 (los costes fijos).

- Así, el modelo del PL es:
  Max $z=250.000*x_1+210.000*x_2$
  $7x_1+5x_2 \leq 175.000$
  $x_1+2x_2 \leq 33.000$
  $x_1 \leq 22.500$
  $x_2 \leq 15.000$
  $x_1,x_2 \in Z^{+}$, toman valores no negativos y enteros.

EJERCICIO 3

Estamos diseñando la producción de una mezcla alimenticia para peces marinos de acuario. Esta mezcla la haremos utilizando algunos de los cinco alimentos siguientes: A_1, A_2, A_3, A_4 y A_5. La composición en calorías y grasas de cada alimento se muestra en la tabla anexa. El precio de coste por kg de cada alimento se muestra a la última fila de la tabla. Hemos decidido que esta mezcla la comercializaremos en sacos cuyo contenido tiene que estar comprendido entre los 18,5 kg y 19 kg de mezcla. Estamos interesados en decidir cuántos kg hay que utilizar de cada alimento con el fin de obtener, a coste mínimo, sacos de mezcla de manera que el contenido mínimo en calorías y grasas del saco de mezcla sea, respectivamente, de 50 y 75.

	A_1	A_2	A_3	A_4	A_5
Calorías/kg	2	4	2	1	3
Grasas/kg	3	1	2	3	5
Precio coste (€/kg)	0,54	0,42	0,36	0,42	0,78

Resolución:

- Definimos las variables siguientes: A_i= kg de alimento Ai en cada saco.
- Restricciones:
  - Calorías: $2A_1+4A_2+2A_3+A_4+3A_5 \geq 50$
  - Grasas: $3A_1+A_2+2A_3+3A_4+5A_5 \geq 75$
  - Cantidad de mezcla: $A_1+A_2+A_3+A_4+A_5 \geq 18{,}5$ y $A_1+A_2+A_3+A_4+A_5 \leq 19$
  - De signo: $A_i \geq 0$. Las variables toman valores reales y no negativos.
- Función objetivo: si se pretende minimizar el coste de producción de cada saco, entonces hay que minimizar $z=0{,}54A_1+0{,}42A_2+0{,}36A_3+0{,}42A_4+0{,}78A_5$.
- Modelo PL del problema:
  Min $z=0{,}54A_1+0{,}42A_2+0{,}36A_3+0{,}42A_4+0{,}78A_5$
  $2A_1+4A_2+2A_3+A_4+3A_5 \geq 50$
  $3A_1+A_2+2A_3+3A_4+5A_5 \geq 75$
  $A_1+A_2+A_3+A_4+A_5 \geq 18{,}5$
  $A_1+A_2+A_3+A_4+A_5 \leq 19$
  $A_i \geq 0$

EJERCICIO 4

El gestor de un área de servicio de una autopista está organizando la distribución del personal de la cafetería para poder cubrir el servicio. El gestor considera que cada día de la semana (del lunes al domingo) está partido en dos franjas horarias: la "diurna", de 6 h a 18 h, y la "nocturna", de 18 h a 6 h. Así, por ejemplo, la franja nocturna del miércoles empieza a las 18 h del miércoles y acaba a las 6 h del jueves. Las necesidades mínimas de personal durante cada franja horaria de 12 horas en que están divididos los días se recogen en la tabla anexa. El gestor contempla las especificaciones siguientes:

- Para ser rentable el negocio, el número total de trabajadores contratados no tiene que superar las 60 personas.
- Hay tres tipos de trabajadores: "diurnos", "nocturnos" y "domingos". Cada trabajador sólo puede ser de un tipo.
- El día que trabaje, el trabajador "diurno" y "nocturno" hará 12 h seguidas de servicio (toda la franja). El trabajador "domingo" entrará a trabajar el sábado a las 18 h y terminará el lunes a las 6 h. [Nota: El gestor es consciente de que el trabajador tipo "domingo" hace 36 horas seguidas, pero no le preocupa porque ha dado instrucciones para que entre ellos se combinen horas de descanso. Además, el gestor ha "inflado" las necesidades con el fin de que "sobren" trabajadores y puedan, de esta manera, repartir ratos de descanso.]
- Un trabajador que sea tipo "diurno", o uno que sea "nocturno", trabaja tres días a la semana. El gestor puede asignar a cada trabajador uno de los cuatro "turnos" diferentes que existen: el turno de L-Mi-J, el de M-J-V, el de Mi-V-S, y el de S-L-M.
- El salario de los trabajadores es diferente según el tipo al que pertenecen. Si se toma como referencia el salario de los trabajadores "diurnos", los "nocturnos" cobran un 10% más que los "diurnos", y los tipos "domingos" un 25% más que los "diurnos".

El director de personal de la empresa necesita saber a cuántos trabajadores tiene que contratar de cada tipo y cómo tiene que repartirlos de manera que las necesidades de personal queden cubiertas. Lógicamente, el gestor quiere organizar el servicio de forma que el coste sea el mínimo.

Día	L.	M.	Mi.	J.	V.	S.	D.
Diurna	15	16	18	20	22	27	14
Nocturna	6	5	7	7	16	24	14

Resolución:

- Variables: llamamos D_1, D_2, D_3, y D_4 al número de trabajadores "diurnos" de cada turno, respectivamente; llamamos N_1, N_2, N_3, y N_4 al número de trabajadores "nocturnos" de cada turno; llamamos G al número de trabajadores tipo "domingos".
- Restricción de rentabilidad: $D_1+...+D_4+N_1+...+N_4+G\leq 60$;
- Restricciones de necesidad de personal:
  - Lunes: $D_1+D_4\geq 15$; $N_1+N_4\geq 6$;
  - Martes: $D_2+D_4\geq 16$; $N_2+N_4\geq 5$;
  - Miércoles: $D_1+D_3\geq 18$; $N_1+N_3\geq 7$;
  - Jueves: $D_1+D_2\geq 20$; $N_1+N_2\geq 7$;
  - Viernes: $D_2+D_3\geq 22$; $N_2+N_3\geq 16$;
  - Sábado: $D_3+D_4\geq 27$; $N_3+N_4+G\geq 24$;
  - Domingo: $G\geq 14$;
- Restricciones de signo: todas las variables son enteras no negativas. Es un problema de PL.
- Función objetivo: hay que minimizar el coste de los trabajadores contratados teniendo en cuenta que el coste de un N_k es 1,1 veces el de un D_k, y que el coste de un G es 1,25 veces el de un D_k. Es decir, hay que minimizar la función $z=D_1+...+D_4+1.1N_1+...+1,1N_4+1,25G$.
  - El modelo del PL es:
    Min $z=D_1+...+D_4+1.1N_1+...+1.1N_4+1.25G$

$D_1+...+D_4+N_1+...+N_4+G \leq 60$
$D_1+D_4 \geq 15$; $N_1+N_4 \geq 6$
$D_2+D_4 \geq 16$; $N_2+N_4 \geq 5$
$D_1+D_3 \geq 18$; $N_1+N_3 \geq 7$
$D_2+D_3 \geq 22$; $N_2+N_3 \geq 16$
$D_3+D_4 \geq 27$; $N_3+N_4+G \geq 24$
$G \geq 14$
$D_k, N_k, G \in Z^+$

EJERCICIO 5
Suponemos que, a partir de una situación real, hemos llegado a la modelización del siguiente modelo de PL:

$$
\begin{array}{ll}
[\text{MAX}]\ Z = 3X+Y & \\
(1) & X + Y \geq 3 \\
(2) & -2X + Y \leq 3 \\
(3) & 4X + Y \leq 9 \\
 & X,Y \geq 0.
\end{array}
$$

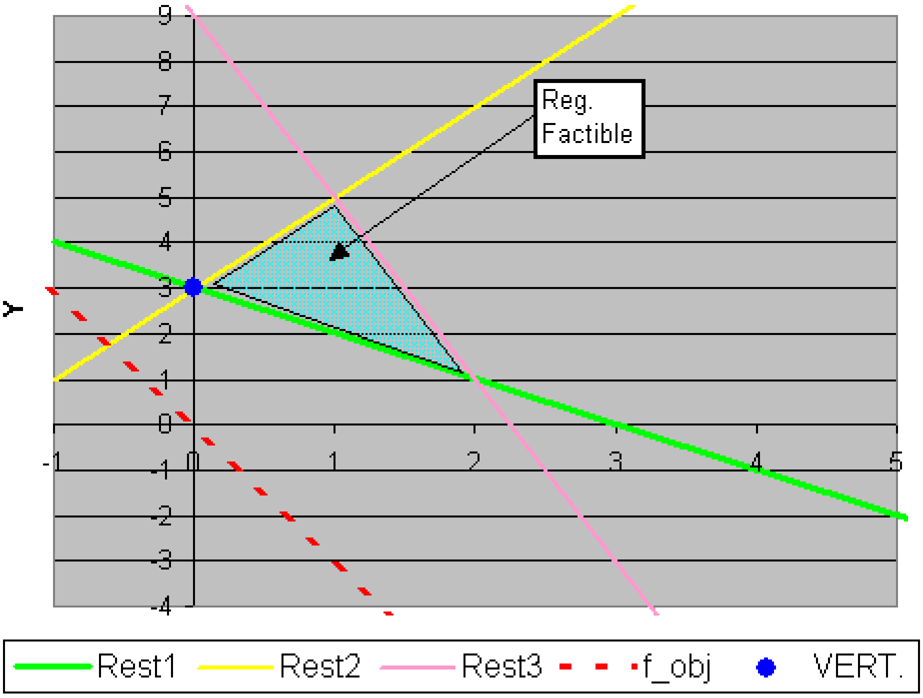
Utilizad el programa Excel y ayudaos del método de resolución de los vértices y del método de la línea de isobeneficio para contestar las cuestiones siguientes:

a) Representad la región factible y determinad los vértices que la delimitan.
b) Determinad el vértice óptimo y el valor máximo de la función objetivo. ¿De qué tipo de solución se trata?
c) Determinad y describid cuál sería la solución óptima si la función objetivo fuera [MAX] Z = 8X+2Y.
d) Determinad y describid cuál sería la solución óptima si la función objetivo fuera [MIN] Z = 3X+Y.

Resolución:

En la resolución de este ejercicio se ha utilizado el archivo "**Módulo 1_Guia_Ejer_5.xls"** que hemos adjuntado a esta guía.

a) Representamos la región mediante el programa Excel. Sustituyendo el punto (X,Y)=(0,0) en cada una de las restricciones deducimos qué región (semiplano) es factible para cada una de las restricciones. Entonces, la región factible del PL es la intersección de los tres semiplanos (ver la hoja **resol_graf** del archivo de Excel **"Módulo 1_Guia_Ejer_5.xls"**):



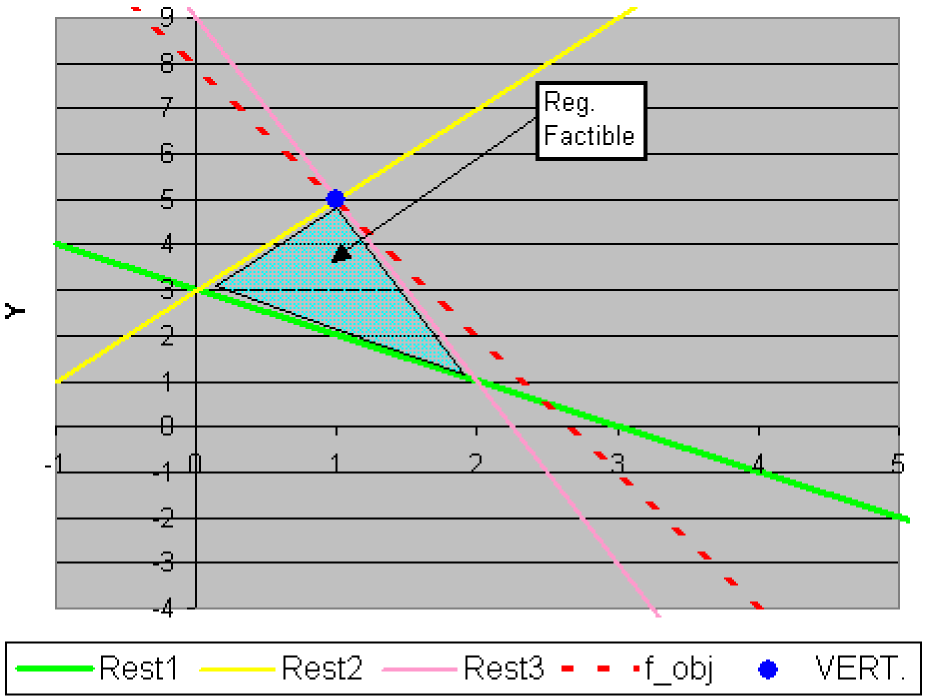
Para calcular los vértices utilizamos el rango de casillas **L11:O20** de la hoja. Por ejemplo, la siguiente tabla, nos permite calcular el vértice intersección de las restricciones rest1 y rest2.

	Coef X	Coef Y	Ter. indep.
rest1	1	1	3
rest2	-2	1	3
F. Obj.	3	1	
VERT.	X	Y	
	0	3	
Z=	3		

Se observa como en el gráfico también se representa gráficamente el vértice determinado. Procediendo de este modo podemos calcular los 3 vértices que delimitan la región:

(0,3) (1,5) (2,1)

b) Usamos el método de la línea isobeneficio para determinar el vértice óptimo, vamos dando valores diferentes a la casilla **Y8** y confirmamos que el vértice donde la Z maximiza es el vértice donde se corta la restricción rest2 con la rest3. Este vértice es el vértice (X,Y)=(1,5) y la función objetivo toma el valor Z=8. Podemos concluir que la solución del PL es una solución única.

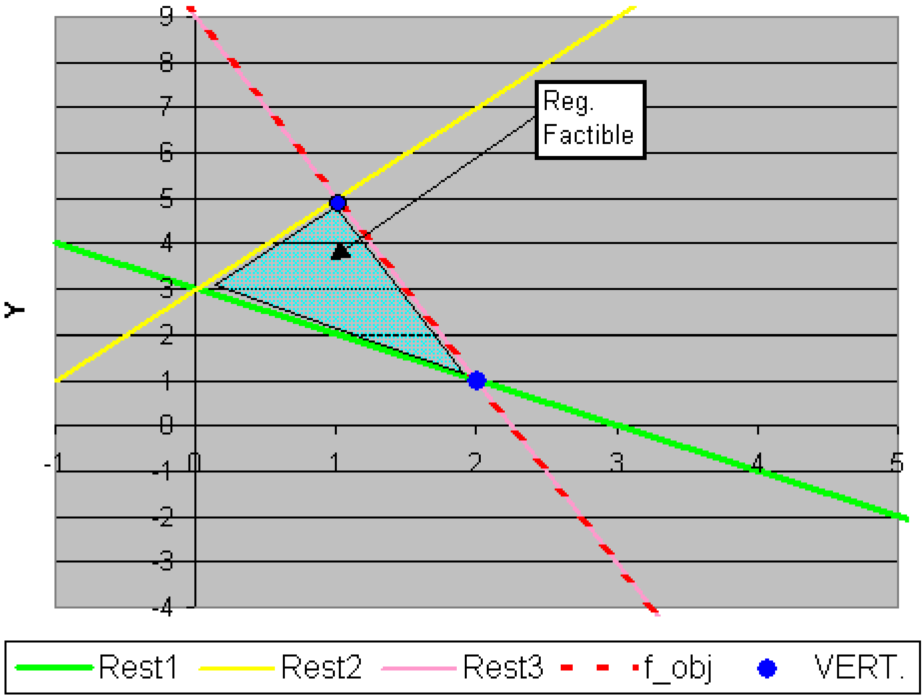


c) Dado que ahora Z=8X+2Y, hace falta modificar las fórmulas de las casillas **L5** y **L6** para cambiar el valor 3 (antiguo coeficiente de la X) por el valor 8 (nuevo coeficiente de la X) y el valor 1 (antiguo coeficiente de la Y) por el valor 2 (nuevo coeficiente de la Y). También hace falta cambiar los valores de las casillas **M14** y **N14** (ver hoja **resol_graf.C**). A efectos prácticos lo que sucede es que la inclinación de la función objetivo ha cambiado. Si nos fijamos en las casillas del cálculo del vértice, cuando hacemos intervenir las restricciones rest2 y rest3 observamos que:

	Coef X	Coef Y	Ter. indep.
rest2	-2	1	3
rest3	4	1	9
F. Obj.	8	2	
VERT.	X	Y	
	1	5	
Z=	18		

ahora, en el vértice (X,Y)=(1,5) la función objetivo toma el valor Z=18. Introducimos este valor 18 en la casilla **I8** y observamos que la función objetivo coincide con todo un segmento de la región factible. El otro vértice de este segmento es el vértice determinado por las restricciones rest1 y rest3. Por lo tanto, este vértice (X,Y)=(2,1) también será óptimo:

	Coef X	Coef Y	Ter. indep.
rest1	1	1	3
rest3	4	1	9
F. Obj.	8	2	
VERT.	X	Y	
	2	1	
Z=	18		



Podemos concluir que la solución del PL es una solución múltiple acotada y el segmento solución es:

$$\alpha(1,5)+(1-\alpha)(2,1), \text{ donde } \quad 0\leq\alpha\leq 1 \text{ y } Z=18.$$

En la hoja **resol_graf.C2**, se pueden obtener los diferentes puntos del segmento dando valores entre 0 y 1 en la celda **M28.**

d) De la misma forma, usamos el método de la línea isobeneficio para determinar el vértice óptimo; vamos dando valores diferentes a la casilla **I8** y confirmamos que el vértice donde Z alcanza el mínimo es el vértice donde se corta la restricción rest1 con la rest2. Este vértice es el punto (X,Y)=(0,3) y la función objetivo toma el valor Z=3. Podemos concluir que la solución del PL es una solución única.