HEURÍSTICA Y OPTIMIZACIÓN

PRÁCTICA 1

Programación Lineal

https://github.com/100432039/proyecto1_heuristica_100432147_100432039

Jaime Sanz Fraile	100432147
Ignacio Marin de la Barcena Perez	100432039

ÍNDICE

INTRODUCCIÓN

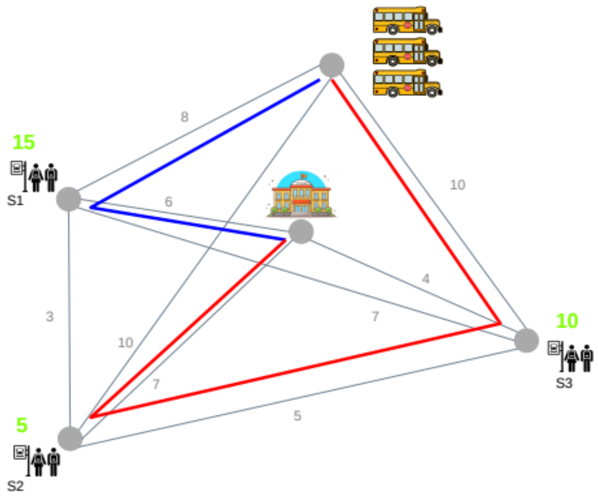
El objetivo de esta práctica es aprender a modelar problemas de programación lineal y a resolver estos mediante el uso de dos herramientas diferentes: hojas de cálculo y algoritmos de resolución usando técnicas de modelización.

Para ello, se plantean varios problemas relacionados con el diseño de rutas los cuales habrá que modelar y resolver mediante las herramientas previamente nombradas. En primer lugar, se modelará un problema básico de rutas que servirá de base para la posterior introducción de ciertas restricciones complicando la modelización de este y la necesidad de otra herramienta para su correcta resolución.

Por último, se analizarán los datos obtenidos, así como el correcto funcionamiento de las restricciones impuestas en el modelo mediante el cambio de los datos. Además, se discutirán las ventajas y desventajas que existen entre las diferentes herramientas utilizadas en esta práctica.

MODELO BÁSICO

Antes de empezar queremos aclarar que las filas de todas las matrices que aparecen en esta práctica representan la parada desde la que sale el bus y las columnas la parada de destino. El primer problema que nos plantean es la generación de las rutas que los autobuses deben seguir para recoger a los alumnos en sus paradas y llevarlos al colegio.



Se pide determinar las rutas de forma que se minimice el coste del transporte escolar. Para ello, se nos indica que por cada autobús utilizado existen unos costes fijos de 120€, además por cada km recorrido existirán costes variables equivalentes a 5€. Obtenemos así la siguiente función objetivo:

$$min\ z = 120 * \sum_{j=1}^{5} X_{ij} + 5 * \sum_{i=1}^{5}\sum_{j=1}^{5} X_{ij} * C_{ij}$$

Para que esta función y sus respectivas variables tengan sentido habrá que definir las variables de decisión del problema así como presentar los datos de este. Comenzaremos presentando los datos:

- CA → nos indica la capacidad máxima de los autobuses = 20
- R → número máximo de rutas disponibles = 3
- C_{ij} → Matriz de costes, representa la distancia en km desde una parada a otra.
  - $C \in N$
  - $i, j \in \{1,\ p\}$ → p indica el número de paradas

	p	s1	s2	s3	c
p	∞	8	10	10	∞
s1	∞	∞	3	7	6
s2	∞	3	∞	5	7
s3	∞	7	5	∞	4
c	∞	∞	∞	∞	∞

- A_j → Vector de alumnos, representa los alumnos que hay en cada parada al comienzo del problema.
  - $A \in N$
  - $j \in \{1, p\}$ → p indica el número de paradas

	p	s1	s2	s3	c
p	0	15	5	10	0

Las variables de decisión de este problema son dos matrices, la primera matriz (X) es una matriz binaria que nos indicará la ruta óptima que seguirán los autobuses para ir desde el parking al colegio recogiendo a todos los alumnos. La segunda matriz (Y), nos indica el flujo de alumnos de una parada a otra dada la ruta o rutas óptimas en la matriz X, en concreto, cada casilla representará el número de alumnos que llegan en el autobús desde la parada que indique la fila a la parada destino que indique la columna.

- X_{ij} → Matriz de rutas.
  - $X \in \{0, 1\}$
  - $i, j \in \{1, p\}$ → p indica el número de paradas

	p	s1	s2	s3	c
p	0	1	0	1	0
s1	0	0	1	0	0
s2	0	0	0	0	1
s3	0	0	0	0	1
c	0	0	0	0	0

- Y_{ij} → Matriz auxiliar, indica el flujo de los alumnos que llegan a cada parada.
  - $Y \in N$
  - $i, j \in \{1, p\}$ → p indica el número de paradas

	p	s1	s2	s3	c
p	0	0	0	0	0
s1	0	0	15	0	0
s2	0	0	0	0	20
s3	0	0	0	0	10
c	0	0	0	0	0

Para que estas variables de decisión no puedan tomar cualquier valor, el problema presenta una serie de restricciones que se tienen que cumplir. En concreto encontramos 9 restricciones para la matriz X y 29 para la matriz Y. Para la matriz X encontramos las siguientes restricciones:

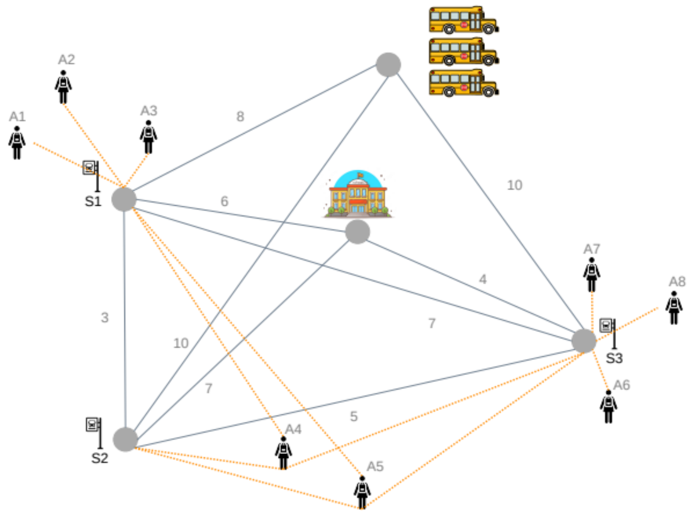
- $\sum_{j=1}^{5} X_{1j} \leq 3$ → comprobamos que no salen más rutas de las disponibles.
- $\sum_{j=1}^{5} X_{1j} = \sum_{i=1}^{5} X_{i5}$ → comprobamos que las rutas que salen del parking llegan al colegio.
- Las siguientes tres restricciones comprueban que únicamente llega un bus a cada una de las paradas con alumnos:
  - $\sum_{i=1}^{5} X_{i2} = 1$ → se encarga de controlar la parada s1.
  - $\sum_{i=1}^{5} X_{i3} = 1$ → se encarga de controlar la parada s2.
  - $\sum_{i=1}^{5} X_{i4} = 1$ → se encarga de controlar la parada s3.
- Las siguientes tres restricciones comprueban que únicamente sale un autobús desde cada una de las paradas con alumnos:
  - $\sum_{j=1}^{5} X_{2j} = 1$→ se encarga de controlar la parada s1.
  - $\sum_{j=1}^{5} X_{3j} = 1$→ se encarga de controlar la parada s2.
  - $\sum_{j=1}^{5} X_{4j} = 1$→ se encarga de controlar la parada s3.
- $X_{ij} = \{0, 1\}$ → comprobamos que cada uno de los valores de esta matriz es binario.

Para la matriz Y encontramos las siguientes restricciones:

- $Y_{ij} \leq CA * X_{ij}$ → comprueba que nunca se exceda la capacidad del autobús.
- Las tres siguientes restricciones se encargan de controlar el flujo de alumnos de cada parada que tenga alumnos esperando, es decir, comprueba que el número de alumnos que sale de una parada es igual al número de alumnos que llega a esa parada más los que se encuentran esperando.
  - $\sum_{j=1}^{5} Y_{2j} = \sum_{i=2}^{4} Y_{i2} + \sum_{i=1}^{5} X_{i2} * A_2$ → se encarga de controlar la parada s1.
  - $\sum_{j=1}^{5} Y_{3j} = \sum_{i=2}^{4} Y_{i3} + \sum_{i=1}^{5} X_{i3} * A_3$ → se encarga de controlar la parada s2.
  - $\sum_{j=1}^{5} Y_{4j} = \sum_{i=2}^{4} Y_{i4} + \sum_{i=1}^{5} X_{i4} * A_4$ → se encarga de controlar la parada s3.
- $Y_{ij} = N$ → comprobamos que cada uno de los valores de esta matriz es natural.

MODELO AVANZADO

Para este segundo ejercicio la administración del colegio se está planteando si podría reducir aún más el coste distribuyendo los alumnos en las paradas de otra manera. Para ello han calculado las distintas paradas a las que podría acudir cada alumno, teniendo en cuenta que la distancia entre su casa y la parada a la que vaya no puede ser mayor que una determinada.



De esta manera, los alumnos A4 y A5 podrían ir a cualquiera de las 3 paradas. Además el colegio tiene información de las familias sabiendo si ciertos alumnos son hermanos entre sí, en cuyo caso deben acudir a la misma parada.

El problema sigue manteniendo el objetivo de minimizar el coste de las rutas lo máximo posible, por tanto, la función objetivo será la misma que en el apartado anterior:

$$min\, z = 120 * \sum_{j=1}^{5} X_{ij} + 5 * \sum_{i=1}^{5}\sum_{j=1}^{5} X_{ij} * C_{ij}$$

Aunque se tenga la misma función objetivo que en el primer apartado, este es un problema totalmente diferente ya que habrá que resolver el problema de asignación de alumnos antes de poder realizar cualquier optimización de las rutas. Por ello, para este apartado nos presentan los siguientes datos:

- CA → nos indica la capacidad máxima de los autobuses = 4
- R → número máximo de rutas disponibles = 3
- n → número de alumnos
- H_{ij} → Matriz de hermanos, indica que alumnos son hermanos entre sí.
  - $H \in \{0, 1\}$
  - $i, j \in \{1, n\}$ → n indica el número de alumnos

	A1	A2	A3	A4	A5	A6	A7	A8
A1	1	0	0	0	0	0	0	0
A2	0	1	0	0	0	0	0	0
A3	0	0	1	0	0	0	0	0
A4	0	0	0	1	1	0	0	0
A5	0	0	0	1	1	0	0	0
A6	0	0	0	0	0	1	0	0
A7	0	0	0	0	0	0	1	0
A8	0	0	0	0	0	0	0	1

- S_{ij} → Matriz de alumnos y paradas, indica que alumnos pueden ir a las paradas.
    - $S \in \{0, 1\}$
    - $i \in \{1, n\}$ → n indica el número de alumnos
    - $j \in \{1, s\}$ → s indica el número de paradas donde se pueden subir alumnos.

	s1	s2	s3
A1	1	0	0
A2	1	0	0
A3	1	0	0
A4	1	1	1
A5	1	1	1
A6	0	0	1
A7	0	0	1
A8	0	0	1

- Además de estas matrices de datos contamos evidentemente con la matriz de costes del apartado anterior (C), aunque el vector de alumnos pasará a ser ahora una variable.

Las variables de decisión de este problema son tres matrices más un vector, la dos primeras matrices son las que están definidas en el primer apartado (X, Y). La tercera matriz (Z) indica la asignación de los alumnos a las paradas donde serán recogidos por el autobús. Por último, el vector (A) que previamente eran datos ahora será una variable de decisión que estará directamente relacionado con la matriz Z.

- Z_{ij} → Matriz de asignación de paradas, indica la parada en la que se recogerá a los alumnos.
    - $Z \in \{0, 1\}$
    - $i \in \{1, n\}$ → n indica el número de alumnos
    - $j \in \{1, s\}$ → s indica el número de paradas donde se pueden subir alumnos.

	s1	s2	s3
A1	1	0	0
A2	1	0	0
A3	1	0	0
A4	0	1	0
A5	0	1	0
A6	0	0	1
A7	0	0	1
A8	0	0	1

- A_j → Vector de alumnos, representa los alumnos que hay en cada parada al comienzo de las rutas.
    - $A \in N$
    - $j \in \{1, p\}$ → p indica el número de paradas en las que se recogen alumnos.

	s1	s2	s3
p	3	2	3

Para que estas variables de decisión no puedan tomar cualquier valor, el problema presenta una serie de restricciones que se tienen que cumplir. Para la matriz Y tenemos exactamente las mismas restricciones que en el primer apartado. Para la matriz X encontramos las siguientes modificaciones:

- Las siguientes tres restricciones comprueban que como máximo pasa un autobús por cada parada:
    - $\sum_{i=1}^{5} X_{i2} \leq 1$ → se encarga de controlar la parada s1.
    - $\sum_{i=1}^{5} X_{i3} \leq 1$ → se encarga de controlar la parada s2.
    - $\sum_{i=1}^{5} X_{i4} \leq 1$ → se encarga de controlar la parada s3.

- Las siguientes tres restricciones comprueban que como máximo sale un autobús desde cada una de las paradas con alumnos:
    - $\sum_{j=1}^{5} X_{2j} \leq 1$→ se encarga de controlar la parada s1.
    - $\sum_{j=1}^{5} X_{3j} \leq 1$→ se encarga de controlar la parada s2.
    - $\sum_{j=1}^{5} X_{4j} \leq 1$→ se encarga de controlar la parada s3.
- El resto de restricciones siguen igual.

Estas modificaciones son necesarias ya que ahora existe la posibilidad de que no sea necesario pasar por todas las paradas dependiendo la asignación de los alumnos. Esta asignación depende directamente de las restricciones impuestas para la matriz Z y el vector A. Para la matriz Z encontramos:

- $\sum_{i=1}^{n} Z_{ij} \leq CA$ → el número de alumnos por parada no supera la capacidad del autobús.
- $\sum_{j=1}^{s} Z_{ij} = 1$ → cada alumno está asignado a una parada.
- $Z_{ij} \leq S_{ij}$ → cada alumno está asignado en una parada en la que puede estar.
- $\sum_{i=1}^{n} H_{ik} * Z_{ij} = Z_{kj} * \sum_{l=1}^{n} H_{kl}$ → comprueba que los hermanos vayan a la misma parada.

Para el vector A encontramos la siguiente restricción:

- $\sum_{i=1}^{n} Z_{ij} = A_j$ → comprueba el número de alumnos por parada.

ANÁLISIS DE RESULTADOS

Una vez modelados ambas partes procedemos a implementar el problema en un lenguaje sofisticado, Mathprog. La implementación se basa en la creación de dos archivos, uno denominado ***data.dat*** y otro ***model.mod***.
En la implementación del primer problema, en el fichero ***.dat*** se especificarán todos los parámetros del problema, aquellos valores que definirán el resultado óptimo del problema. En este caso encontrándose los valores correspondientes a los datos de los costes, el número de disponible de autobuses, la capacidad de los mismos, la matriz de costes y el vector de alumnos por parada (parámetros que cambiaremos más adelante para las distintas pruebas). Así mismo en el fichero ***.mod*** se especificarán todas las variables, restricciones y función objetivo del problema. Los campos que se indican se implementan de forma general con el objetivo de que el problema al ser modificado siempre de la solución óptima.
Para la segunda parte, tanto el fichero ***.dat*** como el ***.mod*** se han modificado para que cumplan con los requisitos de este segundo problema de mayor complejidad. Para ello se han añadido nuevos parámetros (matriz de hermanos y de paradas posibles de cada alumnos) así como se han modificado los parámetros de la parte 1 (vector de alumnos ahora es variable, mirar el punto anterior). Por otro lado, se han añadido nuevas restricciones así como modificado algunas de las anteriores (mirar punto anterior).
Una vez implementados ambos archivos y ejecutados (glpk online o en glpk linux) vamos a analizar los resultados obtenidos para ambas partes:

- Para la parte 1 obtenemos que el coste mínimo es de 400 euros tras recorrer todas las paradas y sacar dos autobuses. Observando la matriz de rutas se extrae que dos rutas son necesarias, una que pase por las paradas 1 y 2 y otra por la 3.Esto confirma que la matriz de rutas obtenida va en sintonía con las restricciones implementadas. No obstante, para confirmar que hace aplica las restricciones especificadas hay que mirar la matriz de flujo obtenida en la que al mirar los valores de la misma y esta cuadrar con las restricciones se puede decir que el programa da la solución óptima para el problema planteado siguiendo las restricciones impuestas y limitada por las restricciones de la capacidad de los autobuses, el control de flujo de los alumnos (que los alumnos que salgan de una parada sea igual a los que llegan más los que se suben) y que todos los buses han de salir del parking y que todas las paradas han de ser visitadas una vez (esto se extrae del fichero de salida ***output.txt***).

- En el segundo problema, el resultado óptimo obtenido es 585 euros, tras recorrer todas las paradas y sacar a los 3 autobuses. Esto se extrae al igual que en el ejercicio anterior de la matriz de rutas dónde especifica que las tres paradas son visitadas por un bus diferente cada una al tener dos alumnos que son hermanos y que tienen que ir juntos a la misma parada, dato que se ve en la matriz variable de asignación. Estos resultados al examinarlos, comprueban que se cumplen con todas las restricciones ejecutadas. Llevando así a la solución óptima del problema general, al asignar correctamente las paradas a las que puede ir cada alumno y el número de rutas necesarias. Por el contrario, en este problema, las restricciones que limitan el problema son otras como se puede observar en el fichero ***output.txt*** son las siguientes, si dos alumnos son hermanos o no, el número máximo de alumnos por parada (en función de la capacidad del autobús) y que no es necesario que se recorran todas las paradas y que cada alumno tenga asignada una parada para así no perder ningún alumno.

Una vez analizados los resultados óptimos para los dos problemas dados, vamos a estudiar la complejidad de las variables y restricciones definidas. En ambos casos (problemas), las restricciones definidas siguen una complejidad lineal al ser del tipo parámetro * variable, sumatorio de variables o suma producto de variables por parámetros (ver ambos modelos). A su vez, las variables implementadas también son de complejidad lineal ya que a pesar de aparecer a lo largo del problema siempre son de grado 1 y no afectan a las otras variables por medio de otro operador que no sea el de suma.

Tras detallar los resultados obtenidos de ambos problemas específicos, procedemos a realizar una serie de pruebas sobre ambos problemas cambiando tantos los distintos parámetros como las variables:

- Pruebas problema 1:

  → Modificación capacidad de los autobuses 20 -> 30:

    El resultado óptimo obtenido es un gasto de 220 euros, dónde al tener cada autobús una capacidad para 30 alumnos solo es necesario que salga un autobús que recorra las 3 paradas y llegue al colegio.

  → Modificación de la matriz de coste por trayecto solo en una dirección, no en ambas. Es decir, el valor de ir de la primera parada a la segunda, no tiene porque ser el mismo de ir de la segunda parada a la primera:

    El resultado obtenido es de un coste de 320 euros, dónde los autobuses siguen el mismo formato de rutas que el problema original pero en este caso es el bus que va a la tercera parada el que luego va a la segunda parada.

  → Modificación capacidad de los autobuses 20 -> 10, matriz de costes y el número de alumnos en cada parada pasa a ser 5:

    Para esta prueba obtenemos un valor de coste de 310 euros, dónde las rutas salen a la parada 1 y 3. Y al igual que en la prueba anterior es el bus que va a la parada 3 el que va a recoger a los alumnos de la parada 2.

  → Aumentamos el número de paradas 3 -> 4, resto de parámetros iguales:

    En esta última prueba para el problema 1 obtenemos un resultado de que el mínimo coste es de 405 euros. En está solución observamos que salen dos autobuses, uno primero a la parada 2 y otro a la parada 4, dónde el que va a la parada 2 va a la parada 3 y de ahí al colegio y el autobús que va a la parada 4 va después a la parada 1 y al colegio. Al examinar el fichero output.txt se ve que está modificación del problema original cumple con todas las restricciones del problema dando a suponer que da el resultado óptimo.

Nota*: No se han realizado pruebas para los parámetros del coste de un autobús y coste por kilómetro recorrido, al considerarse que la modificación de estos parámetros y dejar el resto con el mismo valor solo aumentará el valor o disminuirá el coste pero no se modificará la matriz de variables.

- Pruebas problema 2:
  → Modificación de la matriz de hermanos (no hay hermanos):

  El resultado obtenido es 380 euros, al no existir hermanos, los alumnos A4 y A5 son asignados cada uno de ellos a las paradas 1 y 3 respectivamente siendo la diferencia con el problema 2 original en que en este caso solo son necesarios 2 autobuses en vez de tres reduciendo en gran escala el coste además de cumplir con todas las restricciones implementadas.

  → Modificación de la matriz de hermanos y modificación de las posibilidades de paradas a las que puede ir cada alumnos:

  El coste mínimo es de 380 euros. En este caso, el trío de hermanos (A4, A5 Y A7), son asignados a la parada 3 dónde con el alumno A8 cumplen todas las restricciones implementadas, por otro lado, los alumnos A1 y A2 (son hermanos) son asignados a la misma parada, en este caso la parada 1 con los alumnos A3 y A6 obteniéndose así la solución óptima.

  → Modificación del problema añadiendo más alumnos (A9, A10, A11 y A12) y paradas (parada_s4).

  Para está última prueba, el resultado óptimo es de 540 euros. Obtenido al asignarse 4 alumnos (A1, A2, A3 y A10) a la parada 1, 0 alumnos a la parada 2, 4 alumnos (A4, A5, A6 y A7) a la parada 3 y 4 alumnos (A8, A9, A11 y A12) a la parada 4. Siendo necesario los tres autobuses para recoger a cada alumnos y no siendo necesario pasar por todas las paradas.

Nota*: No se realizarán pruebas a la modificación de los parámetros ya probados en el problema 1 al ser el problema dos una extensión y modificación del problema 1.

Finalmente, comentaremos las ventajas y desventajas apreciadas en las plataformas office (CALC) y en Mathprog (GLPK). Respecto a la sencillez y visibilidad ***"calc"*** nos ha parecido mucho más fácil de implementar y visualizar. No obstante, a la hora de implementar las restricciones, y las variables se emplea un mayor tiempo con respecto a su implementación en ***"glpk"***, dónde se pueden implementar un gran número de restricciones en una sola línea de código. En relación a la complejidad de los problemas que se pueden implementar en ambas plataformas y el posible alcance como análisis que los resultados pueden proporcionar consideramos que la implementación en ***"glpk"***, es más precisa, a la par de general permitiendo un análisis más detallado de los problemas como la posibilidad de hacer grandes variaciones de los mismos solo teniendo que cambiar los datos del problema y no el fichero en su totalidad.

CONCLUSIÓN

A lo largo de este proyecto hemos consolidado todos los conocimientos adquiridos en la asignatura permitiéndonos ver una aplicación en el mundo real. A su vez, hemos aprendido a usar dos nuevas herramientas con las que resolver ejercicios de programación lineal que esperamos poder utilizar durante nuestra etapa laboral.

Aun así, queremos recalcar, las grandes dificultades que nos hemos encontrado tanto a la hora de modelar ambos ejercicios como en su posterior implementación en las tecnologías previamente nombradas.