

Localización de Servicios

García Prado, Sergio

14 de marzo de 2017

Resumen

Problemas de Localización de servicios <https://github.com/garciparedes/mosel-examples/tree/master/service-location-examples> [2]

1. SET COVERING PROBLEM: DISTANCIAS

El problema de *set covering* o *cubrimiento de conjuntos* consiste en la asignación de un conjunto de recursos n recursos x_j cuyo uso tiene un coste de c_j para cumplir m necesidades. Las necesidades que cubre cada recurso se representan a través de a_{ij} . La modelización matemática de este problema se muestra en la ecuación (1).

$$\begin{aligned} \text{Minimizar} \quad & \sum_{j=1}^n c_j x_j \\ \text{sujeto a} \quad & \sum_{j=1}^n a_{ij} x_j \geq 1, \quad i = 1, \dots, m \\ & x_j \in \{0, 1\}, \quad j = 1, \dots, n \end{aligned} \tag{1}$$

Equación 1: *Formulación del Problema de Cubrimiento Total.*

1.1 Distancia de cubrimiento inferior a 50km

En este caso se ha propuesto resolver el problema de cubrimiento máximo para la asignación de los lugares donde situar los puntos de servicio de entre 30 ciudades (por tanto $m = n = 30$) para poder abastecer a todos ellos en una distancia inferior a 50 km. Puesto que todos los recursos presentan el mismo coste, en este caso $\forall j \ c_j = 1$.

La solución a este problema se muestra en la ecuación (2), la cual requiere de **18** recursos

$$x_j = \begin{cases} 1 & j \in S = \{2, 4, 5, 9, 10, 11, 12, 13, 14, 16, 17, 18, 21, 22, 27, 28, 29, 30\} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

1.2 Distancia de cubrimiento desde 1 a 250km

En este caso, se ha resuelto el mismo problema que en el apartado anterior, pero esta vez de manera iterativa conforme a las distancias necesarias mínimas de cubrimiento en el intervalo $dc \in [1, 250]$. Dichos resultados se muestran en la figura 1. Los resultados completos de todas las iteraciones pueden obtenerse ejecutando el correspondiente fichero `mosel[2]` y examinando el fichero CSV resultante.

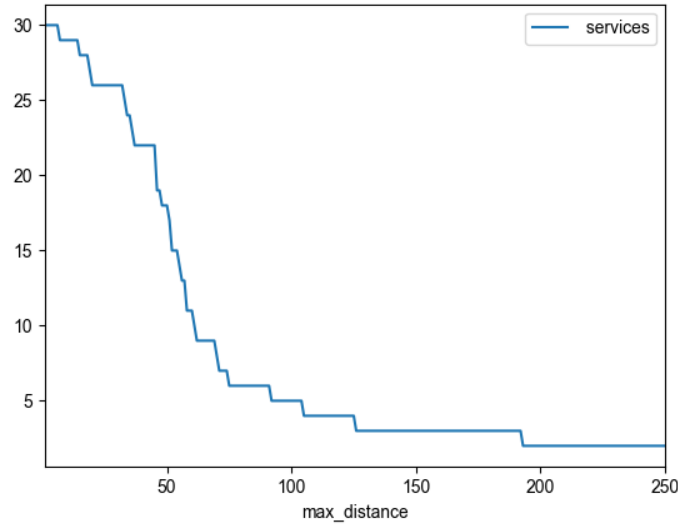


Figura 1: Relación entre la distancia mínima de para que una ciudad se considere cubierta y el número de recursos necesarios para dicha tarea.

2. SET COVERING PROBLEM: DATOS DISPERSOS

En este caso, el problema posee las mismas características que el descrito en la sección 1, cuya modelización matemática se muestra en la ecuación (1). Sin embargo, la novedad que tiene este respecto del anterior es que en este caso los datos son suministrados en forma una matriz dispersa, lo que reduce el tamaño del fichero de datos por lo que se prescinde de las entradas cuyo valor es 0.

2.1 Localización de centros de Ambulancias

En este caso se pide resolver el problema de cubrir $m = 20$ distritos a partir de $n = 10$ centros de ambulancias. Para conocer en qué distritos son cubiertos por qué puntos de servicio se suministra además la matriz dispersa a codificada tal y como se indica en la (3). Tal y como ocurre en el apartado anterior, los costes también son constantes por lo que $\forall j \ c_j = 1$

$$a_{ij} = \begin{cases} 1 & \text{El distrito } i \text{ es cubierto por el centro de ambulancias } j \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Los puntos de asignación óptimos para este problema se muestran en la ecuación (4), es decir, para cubrir las necesidades de todos los distritos se han de colocar 6 centros de ambulancias.

$$x_j = \begin{cases} 1 & j \in S = \{2, 3, 4, 6, 8, 10\} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

2.2 Planificación de Viajes en Empresa de aviación American Airlines

Este problema se basa en la planificación acerca de la tripulación de una empresa de aviación, de tal manera que se aproveche lo más posible la jornada laboral de sus trabajadores. En este

caso, se trata de planificar la tripulación que formará parte de $m = 12$ vuelos, que se compone de un total de $n = 15$ trabajadores. En este caso la codificación de la matriz a sigue la misma distribución que el ejercicio 2.1, por lo que la descripción de la ecuación (3) sigue siendo válida. La principal novedad con respecto a los casos anteriores es que el vector de costes c_j en este caso ya no toma el valor unidad, sino que se refiere al sueldo de cada uno de los trabajadores, es decir, $c_j = \text{Sueldo del trabajador } j$.

La asignación óptima se muestra en la ecuación (5) y presenta un coste total de $9100 = 2900 + 2600 + 3600$.

$$x_j = \begin{cases} 1 & j \in S = \{1, 9, 12\} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

3. SET COVERING PROBLEM: SAYRE-PRIORS

El ejercicio de *Sayre-Priors* presenta el mismo planteamiento que el anterior del apartado 2.2, por tanto la descripción que se ha realizado en el anterior caso es válida para este. Los únicos cambios son los datos de entrada, que en este caso tienen la siguiente dimensionalidad: Se debe encontrar el resultado óptimo para $m = 10$ vuelos y una tripulación de $n = 37$ trabajadores.

La asignación óptima se muestra en la ecuación (6) y genera un coste de $2 + 2 + 3 + 3 = 10$ mil dolares.

$$x_j = \begin{cases} 1 & j \in S = \{12, 24, 29, 32\} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

4. MAX COVERING PROBLEM

En esta sección se trata el *problema de cubrimiento máximo* o *max covering problem*. El problema consiste en lo siguiente: Sea m el número de puntos de demandas y n el de puntos de servicio. El objetivo se trata de maximizar el beneficio h_i obtenido de cubrir el i -ésimo punto de demanda. Para modelizar dicho cubrimiento se utiliza la variable binaria z_i . Para representar los puntos de servicio utilizados se utiliza la variable de tipo binario x_j . La motivación del problema consiste en encontrar el conjunto de variables x_j con cardinalidad máxima denominada por p y prefijada previamente, que maximice la ganancia debida al cubrimiento de los puntos de servicio z_i .

$$\begin{aligned} &\text{Maximizar} && \sum_{i=1}^m h_i z_i \\ &\text{sujeto a} && \sum_{\substack{j \in N_i \\ n}} x_j \geq z_i, && i = 1, \dots, m \\ &&& \sum_{j=1}^n x_j \leq p, \\ &&& x_j \in \{0, 1\}, && j = 1, \dots, n \\ &&& z_i \in \{0, 1\}, && i = 1, \dots, m \end{aligned} \quad (7)$$

Equación 7: Formulación del Problema de Cubrimiento Máximo.

En esta sección se proponen tres ejercicios para resolver este problema, los cuales han sido resueltos para distintos valores de p .

4.1 Usuarios Alcanzados a partir de anuncios en p Revistas

En este problema se trata de encontrar el máximo número de personas cubiertas a partir de la publicación de anuncios en revistas. Los datos de entrada están formados por $n = 10$ revistas que pueden ser alcanzadas por $m = 50$ personas. Para ello se proporciona la matriz a de carácter binario representando la posición a_{ij} que la revista j es leída por la persona i . Se pide resolver el problema para los valores de $p = [1, 10] \in N$. Los resultados se muestran en la tabla 1.

p	% Cubierto	Puntos Abiertos
1	32 %	{5}
2	48 %	{5, 6}
3	64 %	{2, 3, 5}
4	76 %	{3, 5, 6, 7}
5	84 %	{3, 5, 6, 7, 8}
6	90 %	{1, 3, 5, 6, 7, 8}
7	94 %	{1, 3, 5, 6, 7, 8, 10}
8	96 %	{1, 3, 4, 5, 6, 7, 8, 10}
9	98 %	{1, 3, 4, 5, 6, 7, 8, 9, 10}
10	98 %	{1, 2, 3, 4, 5, 6, 7, 8, 9, 10}

Tabla 1: Resultados del ejercicio 4.1.

4.2 Zonas Cubiertas a partir de p de centros de Ambulancias

En este ejercicio se pide resolver el problema de máximo cubrimiento para los valores de $p = [1, 10] \in N$ al igual que en la sección anterior. En este caso el problema se refiere al cubrimiento de $m = 20$ zonas geográficas a partir de $n = 10$ ambulancias, por tanto se trata de encontrar el máximo número de zonas cubiertas con la utilización de p ambulancias.

En este caso, además de la matriz a que codifica de manera binaria en a_{ij} que la zona i es cubierta por la variable j , se proporciona el vector h , que representa en la componente h_i el beneficio obtenido por cubrir la zona i . Esto podría interpretarse como el número de personas que se abarcan en dicha zona.

Los resultados de en este caso se muestran en la tabla 2. Tras analizar los resultados obtenidos se puede obtener como conclusión que el valor óptimo es $p = 6$, es decir, con la utilización de 6 ambulancias se cubren las 20 zonas.

4.3 Zonas Cubiertas a partir de p centros de Servicio Sanitario

Este ejercicio se refiere a la resolución dos problemas de cubrimiento máximo de gran tamaño. Se pide resolver para los valores de $p = [1, 10] \in N$ en ambos casos. Debido al tamaño del problema, en este caso se considera que una zona de demanda i ha sido cubierta si existe una zona de servicio j a una distancia $dc \leq 25$ kilómetros. Al igual que en el caso anterior, también se proporciona un vector h , que en la posición h_i representa la demanda en la zona i .

p	% Cubierto	Puntos Abiertos	f.objetivo
1	44,18 %	{4}	115,8
2	63,94 %	{4, 6}	156,9
3	77,59 %	{4, 5, 9}	190,4
4	83,63 %	{3, 4, 5, 9}	212,6
5	93,76 %	{3, 4, 6, 8, 10}	230,4
6	100 %	{2, 3, 4, 6, 8, 10}	245,4
7	100 %	{1, 2, 3, 5, 6, 8, 10}	245,4
8	100 %	{1, 2, 3, 4, 5, 6, 8, 10}	245,4
9	100 %	{2, 3, 4, 5, 6, 7, 8, 9, 10}	245,4
10	100 %	{1, 2, 3, 4, 5, 6, 7, 8, 9, 10}	245,4

Tabla 2: Resultados del ejercicio 4.2.

El objetivo a maximizar por tanto es la demanda cubierta total. En este caso, la matriz a no es de carácter binario, sino que se compone de atributos reales, positivos, almacenando la posición a_{ij} la distancia en hectómetros de la zona de demanda i al punto de servicio j . Por lo tanto, para utilizar la misma unidad de medida para todas las variables, se transforma la variable $dc = 25 * 10 = 250$ hectómetros.

En el caso del primer fichero de datos de entrada denominado **aint1.dat** se recogen un total de $m = 356$ puntos de demanda y $n = 22$ puntos de servicio. La solución para los distintos valores de p se recogen en la tabla 3.

p	% Cubierto	Puntos Abiertos
1	48,0194 %	{5}
2	66,2764 %	{5, 6}
3	77,9437 %	{5, 6, 12}
4	85,8574 %	{11, 12, 16, 20}
5	91,7617 %	{5, 6, 12, 17, 19}
6	94,4597 %	{5, 6, 9, 12, 17, 19}
7	96,7204 %	{2, 11, 12, 14, 16, 17, 20}
8	97,9665 %	{2, 3, 6, 9, 12, 17, 19, 20}
9	98,3855 %	{2, 6, 8, 9, 15, 17, 18, 19, 22}
10	98,7804 %	{2, 3, 6, 8, 9, 15, 16, 17, 18, 19}

Tabla 3: Resultados del ejercicio 4.3 para aint1.dat.

En el caso del segundo fichero de datos de entrada denominado **aint5.dat** se recogen un total de $m = 328$ puntos de demanda y $n = 19$ puntos de servicio. La solución para los distintos valores de p se recogen en la tabla 4.

p	% Cubierto	Puntos Abiertos
1	69,514 %	{11}
2	84,8082 %	{4, 11}
3	95,3173 %	{3, 4, 14}
4	98,1185 %	{3, 4, 6, 14}
5	99,0051 %	{2, 3, 4, 6, 14}
6	99,5137 %	{1, 2, 3, 4, 13, 14}
7	99,7671 %	{1, 2, 3, 4, 9, 13, 14}
8	99,9904 %	{1, 2, 5, 6, 7, 12, 13, 17}
9	100 %	{1, 2, 4, 5, 6, 7, 13, 14, 17}
10	100 %	{1, 2, 4, 5, 6, 7, 10, 13, 17, 18}

Tabla 4: Resultados del ejercicio 4.3 para aint5.dat.

5. P-MEDIAN PROBLEM Y P-CENTER PROBLEM

En esta sección se resuelven ejercicios del problema de la p-mediana y el p-centro. Por lo tanto, a continuación se describe el modelo matemático de cada uno de estos problemas, así como una descripción acerca del resultado que pretenden optimizar.

En el caso del problema de la p-mediana, modelizado matemáticamente en la ecuación 8, el objetivo es minimizar la distancia global de cada uno de los p puntos de servicio abiertos al conjunto global de puntos de demanda, de manera que los j puntos de servicio abierto mantengan la menor distancia en promedio a los i puntos de demanda.

En este problema, al igual que en los anteriores, se utiliza un vector de demanda denominado h , que en la componente h_i almacena la demanda necesaria por el punto de demanda i . También existe una matriz de distancias de d , que en la posición d_{ij} recoge la distancia del punto de demanda i al punto de servicio j .

Para resolver este problema, además de las variables de decisión x_j utilizadas en casos anteriores, que representan que el punto de servicio j está activo, se añaden las variables y_{ij} , que representan que el punto de demanda i es servido por el punto de servicio j , lo que conlleva que en esta modelización cada punto de demanda sea servido únicamente por un único servicio.

$$\begin{aligned}
& \text{Minimizar} && \sum_{i=1}^m \sum_{j=1}^n h_i d_{ij} y_{ij} \\
& \text{sujeto a} && \sum_{j=1}^n y_{ij} = 1, && i = 1, \dots, m \\
& && y_{ij} \leq x_j, && i = 1, \dots, m, j = 1, \dots, n \\
& && \sum_{j=1}^n x_j = p, \\
& && x_j \in \{0, 1\}, && j = 1, \dots, n \\
& && y_{ij} \in \{0, 1\}, && i = 1, \dots, m, j = 1, \dots, n
\end{aligned} \tag{8}$$

Ecuación 8: Formulación del Problema de la P-mediana.

El problema del p-centro se modeliza de forma matemática tal y como se representa en la ecuación 9. Tal y como se puede apreciar, este problema es un muy similar al de la p-mediana, pero en este caso, la característica que se desea minimizar es la distancia del punto de demanda cubierto que se encuentra a mayor distancia, nótese que con esta estrategia se consigue una minimización del conjunto global de distancias, pero de una manera más equilibrada y uniforme, dado que en el problema anterior, podían existir distancias muy grandes compensadas por otras muy pequeñas como solución óptima, mientras que en este caso es más complicado que se dé una situación semejante.

$$\begin{aligned}
& \text{Minimizar} && w \\
& \text{sujeto a} && \sum_{j=1}^n y_{ij} = 1, \quad i = 1, \dots, m \\
& && \sum_{j=1}^n d_{ij} y_{ij} \leq w, \quad i = 1, \dots, m \\
& && y_{ij} \leq x_j, \quad i = 1, \dots, m, j = 1, \dots, n \\
& && \sum_{j=1}^n x_j = p, \\
& && x_j \in \{0, 1\}, \quad j = 1, \dots, n \\
& && y_{ij} \in \{0, 1\}, \quad i = 1, \dots, m, j = 1, \dots, n
\end{aligned} \tag{9}$$

Equación 9: *Formulación del Problema del P-Centro.*

5.1 P-Median y P-Center aplicado a asignación de p puntos de servicio en 12 municipios

El ejercicio consiste en la resolución del problema del p-centro y la p-mediana para el mismo conjunto de datos. El conjunto de datos se corresponde con las distancias de $n = m = 12$ municipios, codificadas en la matriz d de tal manera que d_{ij} representa la distancia del municipio i al municipio j . Nótese que esta matriz es simétrica, es decir, $d_{ij} = d_{ji}$. Además, la distancia de un municipio a sí mismo es nula, es decir, $d_{ii} = 0$ o lo que es lo mismo, la diagonal principal esta formada por ceros. Además, la demanda de cada municipio se modeliza a partir del vector h , que almacena en la posición h_i la demanda del municipio i .

Las soluciones a los problemas de la p-mediana y el p-centro se muestran en las tablas 5 y 6 respectivamente, para $p = [1, 12] \in N$.

5.2 Problema de la p-mediana sobre aint1.dat y aint5.dat

En este caso, se pide utilizar los conjuntos de datos previamente utilizados en la sección 4.3, por lo tanto tanto la descripción de los datos como del contexto del problema se ha obviado para evitar la repetición. En este caso, el problema que se pide resolver es el de la p-mediana, es decir, la minimización de la distancia máxima de un punto de demanda i a un punto de servicio j . En este caso se ha resuelto para ambos datos de entrada para $p = 5$ instalaciones.

Debido al tamaño del conjunto de datos de entrada, que tal y como se dijo anteriormente presenta $m = 356$ y $n = 22$ en el caso de *aint1.dat* y $m = 328$ y $n = 19$ en el de *aint5.dat*, el número de variables necesarias para la resolución del problema de la p-mediana es muy elevado, por lo que se ha optado por su resolución en el servidor externo **Neos Server**[3]. Las soluciones a cada ejercicio se muestran respectivamente en las tablas 7 y 8.

p	Máx.	Medio	Puntos Abiertos	Max de P_i
1	48	25,79	{9}	{48, 42, 20, 24, 24, 12, 45, 30, 0, 38, 19, 19}
2	33	17	{7, 9}	{18, 33, 20, 24, 24, 12, 0, 15, 0, 22, 19, 19}
3	30	13,18	{1, 6, 11}	{0, 15, 30, 12, 24, 0, 18, 25, 12, 19, 0, 21}
4	28	10,18	{1, 6, 8, 10}	{0, 15, 28, 12, 12, 0, 15, 0, 12, 0, 19, 22}
5	28	7,8	{1, 6, 8, 10, 12}	{0, 15, 28, 12, 12, 0, 15, 0, 12, 0, 19, 0}
6	28	5,85	{1, 6, 8, 10, 11, 12}	{0, 15, 28, 12, 12, 0, 15, 0, 12, 0, 0, 0}
7	15	4,04	{1, 3, 6, 8, 10, 11, 12}	{0, 15, 0, 12, 12, 0, 15, 0, 12, 0, 0, 0}
8	15	2,87	{1, 3, 4, 6, 8, 10, 11, 12}	{0, 15, 0, 0, 12, 0, 15, 0, 12, 0, 0, 0}
9	15	1,98	{1, 3, 4, 6, 7, 8, 10, 11, 12}	{0, 15, 0, 0, 12, 0, 0, 0, 12, 0, 0, 0}
10	15	1,13	{1, 3, 4, 6, 7, 8, 9, 10, 11, 12}	{0, 15, 0, 0, 12, 0, 0, 0, 0, 0, 0, 0}
11	12	0,32	{1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12}	{0, 0, 0, 0, 12, 0, 0, 0, 0, 0, 0, 0}
12	0	0	{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}	{0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0}

Tabla 5: Resultados del ejercicio 5.1 para el problema del p -mediana.

p	Dist. Máxima	Dist. Promedio	Puntos Abiertos
1	47	28	{5}
2	33	18,33	{6, 7}
3	22	14,41	{3, 7, 12}
4	21	10,92	{1, 4, 8, 11}
5	19	12,08	{1, 3, 5, 9, 10}
6	18	7,08	{1, 5, 6, 10, 11, 12}
7	15	5,5	{1, 3, 6, 8, 10, 11, 12}
8	15	5,5	{2, 3, 6, 8, 9, 10, 11, 12}
9	12	3	{1, 2, 3, 6, 7, 8, 10, 11, 12}
10	12	3	{1, 2, 3, 6, 7, 8, 9, 10, 11, 12}
11	12	1	{1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12}
12	0	0	{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}

Tabla 6: Resultados del ejercicio 5.1 para el problema del p -centro.

5.3 Problema de la p -mediana sobre datos geográficos

Los ejercicios que se resuelven en esta sección tienen como novedad respecto de los anteriores la siguiente característica: En este caso los datos de entrada no se presentan a partir de la matriz de distancias, tal y como sucedía en el resto, sino que se ofrecen las coordenadas x e y de cada localización. Esto hace que el problema permita una mayor versatilidad en el sentido de calcular un mayor número de resultados, pero a la vez añade la complicación de requerir el cálculo de las distancias entre puntos.

Para la tarea de calcular las distancias se ha utilizado la *distancia euclídea* para espacios de 2 dimensiones (x, y) , que se define matemáticamente como $d(p, q) = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}$. Por lo tanto, para la modelización del problema de la p -mediana, es necesario calcular dicha medida para todas las posibles combinaciones de localizaciones, de tal manera que la matriz d sea construida siguiendo la expresión $d_{ij} = d(l_i, l_j)$ donde l_i y l_j representan las coordenadas de las localizaciones i y j respectivamente.

p	Dist. Promedio	Dist. Max	Puntos Abiertos
5	94,6455	406	{1, 2, 11, 17, 22}

Tabla 7: Resultados del ejercicio 5.2.1 del problema de la p -mediana para $p = 5$.

p	Dist. Promedio	Dist. Max	Puntos Abiertos
5	59,2697	353	{1, 3, 6, 8, 12}

Tabla 8: Resultados del ejercicio 5.2.2 del problema de la p -mediana para $p = 5$.

En esta sección hay que resolver dos ejercicios. El primero de ellos consiste en las coordenadas (x, y) de $n = m = 15$ poblaciones, para las cuales se pide resolver el problema de la p -mediana para $p = [1, 5] \in N$. Dichos resultados se muestran en la tabla 9.

p	Dist. Total	Dist. Promedio	Puntos Abiertos
1	1101,53	72,9997	{13}
2	705,127	47,6867	{1, 5}
3	527,847	35,2967	{5, 8, 14}
4	422,303	28,1655	{1, 5, 8, 14}
5	329,953	21,8831	{1, 4, 8, 14, 15}

Tabla 9: Resultados del ejercicio 5.3.1 para el problema de la p -mediana.

El segundo caso a resolver presenta las mismas características y problemática que el anterior, sin embargo, el número de poblaciones es diferente, ascendiendo hasta $n = m = 100$, y requiriéndose un total de $p = [7, 10] \in N$ instalaciones para el problema de la p -mediana. Debido a la complejidad generada por el aumento del tamaño del problema, en este caso el problema se ha resuelto utilizando **Neos Server**[3]. Los resultados en este caso se muestran en la tabla 10.

p	Dist. Total	Dist. Promedio	Puntos Abiertos
7	2510,48	24,9789	{4, 6, 28, 42, 59, 60, 90}
8	2290,24	22,6172	{4, 6, 28, 41, 42, 89, 90, 92}
9	2122,31	21,039	{4, 6, 28, 41, 42, 66, 86, 89, 92}
10	1957,9	19,5009	{4, 15, 21, 28, 41, 42, 66, 86, 89, 92}

Tabla 10: Resultados del ejercicio 5.3.2 para el problema de la p -mediana.

REFERENCIAS

- [1] AGUADO, J. S. Modelos de Investigación Operativa, 2016/17.
- [2] GARCÍA PRADO, S. Mosel Examples. <https://github.com/garciparedes/mosel-examples>.
- [3] NEOS SERVER. Solver for Numerical Optimization. <https://neos-server.org/neos/>.