Modelos de Decisión – Trabajo Práctico 2
Master in Management + Analytics, Escuela de Negocios,
Universidad Torcuato Di Tella

Analytics aplicado a residencias médicas: un enfoque basado en programación matemática

Informe del desarrollo del modelo de programación matemática de la consultora DM Analytics sobre la asignación de residentes a instituciones médicas.

- Alumnos: Mauro Bertini y Lucas Veteikis

- Año: 2023

Bibliografía consultada: Hillier, F. S. (2007). Introduction to Management Science. Mc Graw-Hill.

1. Introducción al problema y la decisión

El problema que viene planteado trata de optimizar una medida de satisfacción global (compuesta a su vez de dos medidas de satisfacción individuales) sobre la **asignación** de residentes médicos en determinados hospitales a nivel nacional. El objetivo principal radica en encontrar la combinación ideal que permita a la organización encontrar el **mayor** índice de satisfacción global entre ambos agentes (residentes y hospitales).

Los datos que vienen dados son los listados completos de residentes, hospitales y sus respectivas listas de preferencia. Estas listas de preferencia mantienen una relación ordinal donde la satisfacción es maximizada mientras mayor sea su posición en la asignación (yendo desde el primer puesto hasta el último de acuerdo con las *n* opciones otorgadas).

La entidad gubernamental ya nos ha proveído de un modelo heurístico diseñado bajo la premisa *First Come, First Served* al revisar vacantes de cada hospital disponible y asignar a los residentes que lo tuvieran en su lista de preferencias hasta cubrir la demanda de cada hospital.

El desafío planteado es encontrar un modelo de programación lineal que compita contra este modelo de referencia y encontrar una solución que mejore la métrica de satisfacción global con la intención de modificar el proceso de asignación actual de la entidad gubernamental.

2. Formulación del modelo

Modelo elegido

Para abordar el problema mencionado en el punto anterior, la consultora propone un modelo de programación lineal **de asignación** para la obtención de la solución óptima del problema, con todos los componentes que el mismo conlleva.

Consideramos que este modelo es el apropiado dado que nuestra tarea principal consiste en asignar personas a tareas para maximizar el beneficio obtenido. Esta definición coincide casualmente con un problema de asignación en programación lineal según la bibliografía otorgada (Hillier, 2007, p. 99).

Adicionalmente, la definición de este problema nos permite trabajar con una matriz de restricciones totalmente definida unimodular (TDU) y vectores (right-hand sides) enteros. Esta característica particular de la matriz de restricciones nos garantiza trabajar con números enteros sin necesariamente plantear un problema de programación lineal entera, por lo que su cálculo a nivel computacional es mucho más ágil.

Variable de decisión

Bibliografía consultada: Hillier, F. S. (2007). Introduction to Management Science. Mc Graw-Hill.

La variable de decisión de este problema en particular tomará forma en una variable dicotómica que tomará dos posibles valores: 0, si un residente no es asignado a un hospital, o 1 si el residente si es asignado a un hospital.
La cantidad de variables de decisión para el problema será de una por cada par residente-hospital. Es decir, será de **n*m** (siendo n la cantidad de residentes y m la cantidad de hospitales).

Función objetivo y restricciones

El modelo contará con una función objetivo que buscará la maximización de las sumas de las satisfacciones tanto de los residentes como de los hospitales, a través de la asignación de residentes a hospitales que maximice esta expresión.

A su vez, contará con una restricción para los residentes y otra para los hospitales del modelo.

$$max \sum_{j=1}^{n} \sum_{i=1}^{m} w_{ji}^{r} + w_{ij}^{h}$$

$$s.t: \sum_{i=1}^{m} xji = 1,\ j = 1,\ ...,\ n$$

$$\sum_{j=1}^{n} xji = p_i,\ i = 1,\ ...,\ m$$

$$xji \geq 0,\ i = 1,\ ...,\ m,\ j = 1,\ ...,\ n$$

Donde:

- n representa la cantidad de residentes
- m representa la cantidad de hospitales
- j representa a cada residente
- i representa a cada hospital
- w_{ji}^{r} representa la satisfacción del residente j por ser asignado al hospital i, que es calculado como:
    - $w_{ji}^{r} = \left\{2 * (|L_j^r| - pos(i, L_j^r)\right\} si\ i \in L_j^r\ o\ w_{ji}^{r} = \{-1\}\ si\ i \notin L_j^r$
      Donde L_j^r representa la lista de preferencias del residente j
- w_{ij}^{h} representa la satisfacción del hospital i por que se le asignó al residente j, que es calculado como:
    - $w_{ij}^{h} = \left\{2 * (|L_i^h| - pos(i, L_i^h)\right\} si\ j \in L_i^h\ o\ w_{ij}^{h} = \{-1\}\ si\ j \notin L_i^h$
      Donde L_i^h representa la lista de preferencias del hospital h

Bibliografía consultada: Hillier, F. S. (2007). Introduction to Management Science. Mc Graw-Hill.

- xji representa la variable de decisión, para cada combinación de residente-hospital toma el valor de 1 si se asigna a ese residente a ese hospital y 0 en caso contrario.

En cuanto a las restricciones, la primera quiere expresar que para cada residente j desde 1 hasta n, la suma de las asignaciones de j a cada posible hospital debe ser exactamente de 1, ya que no queremos que el residente sea asignado a más de un hospital o que no sea asignado a ninguno. Esta restricción afecta a todos los residentes. Esto implica que hay **n cantidad de restricciones** para esta primera definición.
La segunda restricción va referida a la capacidad de hospitales. Para cada hospital, la suma de todos los residentes (desde j=1 hasta n) la suma de las asignaciones representadas por la variable dicotómica debe ser igual a p sub i, que representa la capacidad del hospital i, y esta restricción aplica para todos los hospitales desde la i a la m, lo que significa que hay **m cantidad de restricciones** para los hospitales.

Esto garantiza la siguiente formación de la matriz de restricciones TDU siendo el primer subíndice el relacionado a los residentes y el segundo a los hospitales:

X_{11} X_{12} X_{13} … X_{1}□
X_{21} X_{22} X_{23} … X_{2}□
X_{31} X_{32} X_{33} … X_{3}□
…
X□$_{1}$ X□$_{2}$ X□$_{3}$ … X□□

3. Experimentación

Tras la implementación del modelo en CPLEX, adaptamos al programa para que itere sobre cada instancia de cada tamaño planteado (small, medium y large), para que resuelva el problema con las variaciones de cada instancia y versión, para luego de eso guardar las métricas de interés en un dataframe y exportarlo a excel.
Las métricas calculadas para evaluar nuestra solución versus la greedy fueron:

- **Optimal Global Satisfaction:** Esta métrica mide lo que se intenta maximizar en la función objetivo, es la suma de todas las satisfacciones tanto de hospitales como de residentes.
- **Gap:** Diferencia porcentual de la métrica anterior entre la solución Greedy y la solución LP.
- **Residents Satisfaction:** Suma de satisfacción de los residentes de cada simulación. Al igual que con Optimal Global Satisfaction, se calculó el gap porcentual para esta métrica entre LP y Greedy.
- **Hospitals Satisfaction:** Suma de satisfacción de los residentes de cada simulación. También se calculó el gap porcentual para esta métrica entre LP y Greedy.

Bibliografía consultada: Hillier, F. S. (2007). Introduction to Management Science. Mc Graw-Hill.

- **Mutual Satisfaction Count:** Calcula la satisfacción mutua haciendo una suma de +1 por combinación donde el residente es asignado a un hospital que está presente en su lista de preferencia, y el hospital tiene en su lista de preferencia a dicho residente. Luego, se calculó el gap de esta métrica en particular entre el modelo de la competencia y el desarrollado por nuestra consultora.
- **Avg. Mutual Satisfaction:** Toma el resultado de la contabilización de la métrica de Mutual Satisfaction y lo promedia por la cantidad de vacantes disponibles en cada instancia. Esto nos permite hacer comparaciones más precisas entre instancias de mayor tamaño.

4. Discusión y análisis de resultados

Satisfacción global

Evaluando desde la satisfacción global (la métrica que buscamos maximizar en la función objetivo de nuestro modelo de programación lineal) nuestra solución es superior a la solución Greedy en todas las instancias en promedio.

Instance	Mean Optimal Global Satisfaction Greedy	Mean Optimal Global Satisfaction LP	Mean Gap
A. Small	237,20	243,70	2,78%
B. Medium	9579,30	9888,00	3,23%
C. Large	18654,90	19280,70	3,35%

Desde esta perspectiva, no solo el modelo de programación matemática es superior en todas las instancias al greedy, sino que además la brecha se extiende medida que la escala del problema aumenta, por lo cual, es probable que en un problema de escala real el modelo de LP tenga una performance con una brecha aún mayor.

Satisfacción de residentes

Otra métrica que podría ser interesante para el análisis es la satisfacción exclusivamente de los residentes a la hora de ser asignados a un hospital. Evaluandolo desde este punto de vista, el modelo de programación matemática también es superior al codicioso en promedio en todas las instancias. Al contrario de la métrica anterior, a medida que la escala del problema aumenta la brecha de diferencia entre nuestro modelo y el de la competencia es menor.

Bibliografía consultada: Hillier, F. S. (2007). Introduction to Management Science. Mc Graw-Hill.

Instance	Mean Residents Satisfaction Greedy	Mean Residents Satisfaction LP	Mean Residents Satisfaction Gap
A. Small	151,00	160,20	6,28%
B. Medium	8729,20	9065,30	3,85%
C. Large	17818,50	18348,30	2,98%

Satisfacción de hospitales

Evaluando la satisfacción de los hospitales, el modelo codicioso prioriza la satisfacción de los hospitales cuando el problema se evalúa bajo una escala chica o mediana, sin embargo, se nota una mejoría significativa cuando la escala del problema es mayor, como se puede ver en el gap de la satisfacción de hospitales de la instancia grande.

Instance	Mean Hospitals Satisfaction Greedy	Mean Hospitals Satisfaction LP	Mean Hospitals Satisfaction Gap
A. Small	86,20	83,50	-3,26%
B. Medium	850,10	822,70	-2,79%
C. Large	836,40	932,40	14,09%

Satisfacción mutua

La satisfacción mutua es la única métrica en la que el modelo de programación matemática performa consistentemente peor en todas las instancias bajo evaluación cuando se lo compara con el modelo codicioso. Si bien la diferencia no es muy grande, esto sucede ya que el modelo intenta maximizar la suma de las satisfacciones y, a costa de eso, tiende a “sacrificar” una satisfacción para maximizar la otra y la general.

Aún así, es importante destacar que a medida que la escala del problema aumenta, esta brecha entre la satisfacción mutua se va achicando, por lo cual el modelo matemático es cada vez más eficiente en este sentido y el codicioso menos.

Instance	Mean Mutual Satisfaction (G)	Mean Mutual Satisfaction (LP)	Mean Mutual Satisfaction Gap
A. Small	13,40	12,90	-3,61%
B. Medium	63,50	61,40	-3,24%
C. Large	61,90	60,30	-2,52%

Bibliografía consultada: Hillier, F. S. (2007). Introduction to Management Science. Mc Graw-Hill.

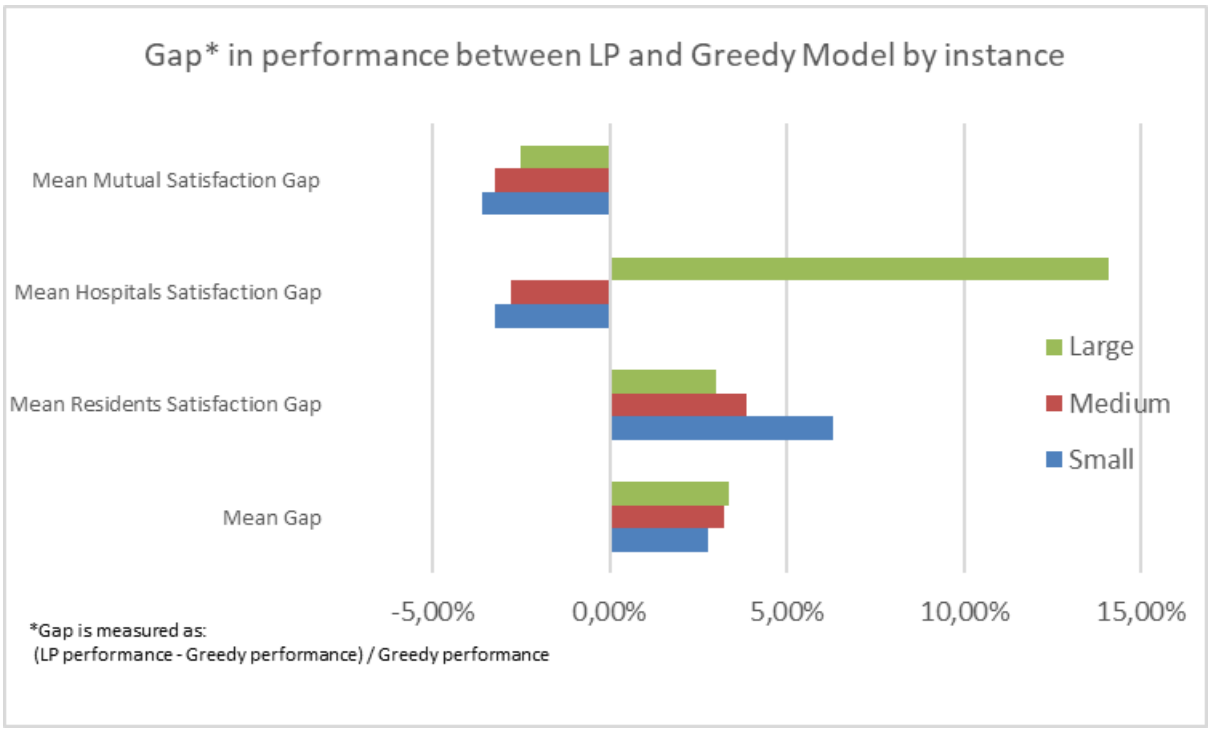
Un enfoque adicional evaluando los promedios de Mutual Satisfaction por instancia y por cantidad de posiciones disponibles nos permite ver la misma tendencia:

Instance	Mean Avg. Mutual Satisfaction (G)	Mean Avg. Mutual Satisfaction (LP)
A. Small	6,7%	6,45%
B. Medium	12,7%	12,28%
C. Large	6,19%	6,03%

En la próxima sección del modelo se sugerirá una posible solución a este aspecto del modelo, que consideramos que es el principal punto débil frente al de la competencia.

Conclusiones del análisis

Como mencionamos en los puntos que recorrimos para el análisis de resultados, el modelo de programación lineal resulta superior en la mayoría de métricas en las que fue probado contra el modelo codicioso de la competencia. La ventaja de nuestro modelo se hace especialmente evidente en escenarios de mayor tamaño, es ahí cuando queda mucho más en evidencia los beneficios de aplicar un enfoque matemático a la resolución del problema. Sin embargo, es importante destacar que se encuentra en la satisfacción mutua un área donde el modelo de programación lineal tiene oportunidad de mejora.



Bibliografía consultada: Hillier, F. S. (2007). Introduction to Management Science. Mc Graw-Hill.

Desde el punto de vista operativo parándose en los zapatos de la entidad gubernamental, podría discutirse el trade-off entre costo y beneficio dado que su solución codiciosa, sin ser óptima, logró un resultado no despreciable.

Quedará en ponderar de cara al cliente si la inversión tiene sentido dados los resultados de performance solamente bajo estas condiciones de evaluación a maximizar. Si se necesitaran refinar las restricciones o modificar condiciones, el modelo de programación lineal daría una herramienta más precisa en ese caso.

5. Limitaciones y posibles extensiones

Como se mostró anteriormente, la métrica de satisfacción mutua es la única que el programa de programación lineal no logra mejorar con respecto a la solución codiciosa planteada. No obstante, dada la flexibilidad que ofrece este tipo de programación, se podría plantear una nueva restricción que permita mejorar el desbalance entre las dos métricas de satisfacción individual siendo esta la siguiente:

$$\sum_{j=1}^{n} \sum_{i=1}^{m} I(j \in res_{\mathrm{i}}) \wedge I(i \in hos\square) / p\square \geq \lambda$$

$$I = (j \in res_{\mathrm{i}}) \wedge I(i \in hos\square) = \mathbf{1}$$

Donde:

- n representa la cantidad total de residentes
- m representa la cantidad total de hospitales.
- I es la función que indica la coincidencia en una lista de preferencias[1].
- $res\square$ es la lista de preferencia de residentes para el hospital i.
- $res\square$ es la lista de preferencia de hospitales para el residente j.
- $\wedge$ representa la unión de ambas funciones como condición
- p representa la cantidad de posiciones vacantes en la instancia.
- s representa la instancia evaluada.
- λ el valor esperado de la proporción de satisfacción mutua que

Esta nueva restricción garantiza un valor mínimo de coincidencia de preferencias entre hospital y residente que permitiría balancear la satisfacción mutua. El valor $\lambda \in \{0,1\}$ garantiza un nivel de discrecionalidad para con la entidad que le permite plantear escenarios con mayor o menor desbalanceo.

[1] En el código, se puede visualizar la condición en la función “get_satisfaction_score” celdas 190 y 191.

Bibliografía consultada: Hillier, F. S. (2007). Introduction to Management Science. Mc Graw-Hill.

Se entiende que un λ que tienda a 1 podría reducir la métrica global de satisfacción en comparación con un modelo que no cuenta con esta restricción e inclusive plantear la infactibilidad. No obstante se prioriza otorgar esta flexibilidad dado que se podría hacer uso de la formulación para objetivos distintos (optimizar la satisfacción global o garantizar una cantidad mínima de actores satisfechos, por ejemplo)

6. Modelo alternativo

Para adaptar el problema planteado a un problema de flujo de costo mínimo se definieron las siguientes condiciones:

Variables de decisión

R_{ji} = Residentes (R_j) a enviar entre cada arco pesado (j,i) que implica la intersección de cada par residente-hospital. Cada residente (R_j) elegirá el arco pesado óptimo que lo vinculará con un único nodo de hospital (H_i).

Los arcos que unen esta combinación tendrán un costo asociado a la métrica de satisfacción conjunta de ambos actores (satisfacción del residente sumado a la satisfacción del hospital acorde a dicha combinación individual).

En el flujo se han diseñado otros arcos cuyo costo es nulo que sirven meramente de transbordo o transición para plantear el problema

Función objetivo

La función objetivo mantiene el esquema del problema de asignación anterior pero modifica su estructura a un problema de flujo de costo mínimo. Las métricas de satisfacción individuales ahora son tomadas como costos en el flujo a minimizar. Para ello, se coloca el signo menos en la formulación para convertir el problema de maximización a minimización:

$$min \sum_{j=1}^{n} \sum_{i=1}^{m} - (w_{ji}^{r} + w_{ij}^{h})$$

Restricciones

En el planteo del grafo conviven 4 tipos de nodos:

1. Nodo 1: un nodo proveedor de residentes cuyo valor por instancia es n.
2. Nodo 2 un nodo destino que recibe las asignaciones de todos los residentes en las posiciones de los hospitales (-n).
3. Nodo r_j: Nodo por cada residente que son recibidos individualmente desde el nodo 1.

Bibliografía consultada: Hillier, F. S. (2007). Introduction to Management Science. Mc Graw-Hill.

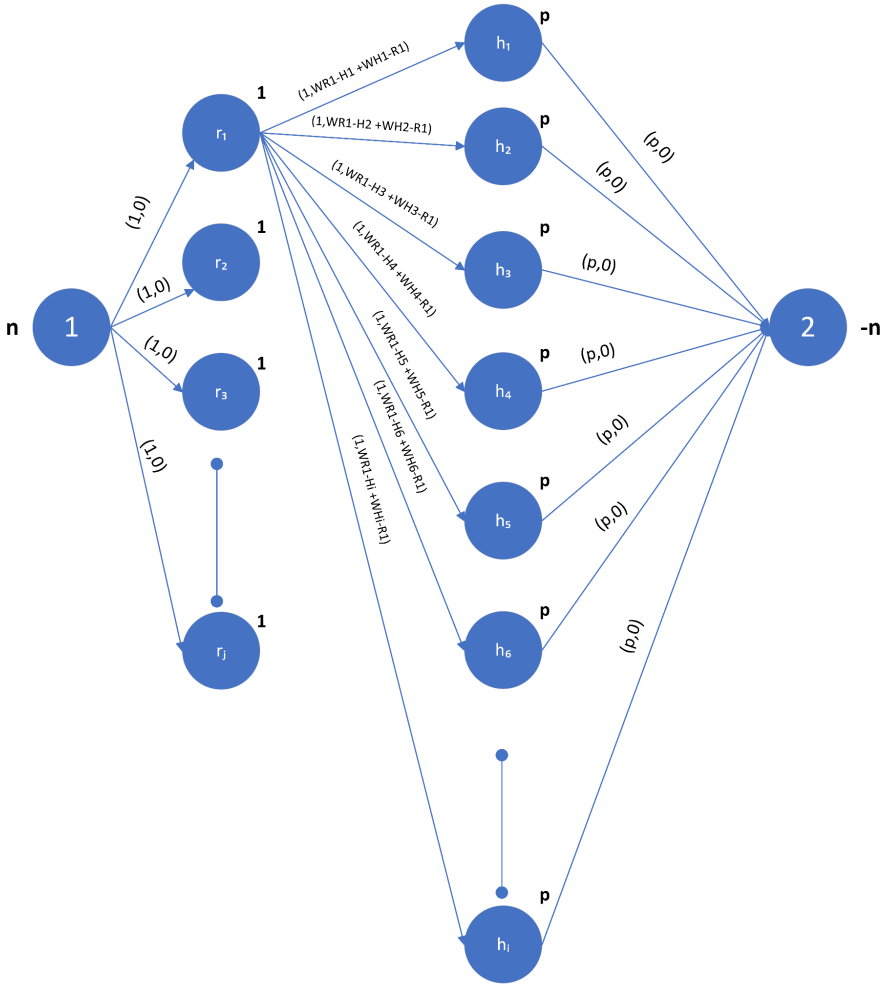
4. Nodo h_i: Nodo para cada hospital cuya demanda es p (cantidad de posiciones requeridas) y recibe p residentes. El resultado de las asignaciones son luego enviadas al nodo 2 donde el flujo finaliza.

Para estos tipos de nodo existen las siguientes restricciones:

- Para nodo 1: $X_{1r1} + X_{1r2} + X_{1r3} + ... + X_{1r□} = n$
- Para cada nodo $r□$: $X_{1r□} - X_{r□□i} = 0$
- Para cada nodo h_i: $X_{r□□i} - X□_{i2} = 0$
- Para nodo 2: $X□_{12} + X□_{22} + X□_{32} + ... + X□_{i2} = -n$

Grafo del problema de flujo de costo mínimo

En el planteo del siguiente grafo, la intención buscada es definir la estructura general y ejemplificar el caso de asignación de un residente. Dado que las instancias tienen distintos tamaños y estos son inabarcables en un gráfico explicativo, se decide recortarlo con la intención de ganar claridad:



En este ejemplo vemos la partición del nodo 1 en *n* nodos *r□* identificados individualmente como los residentes que tienen demanda 1. Enfrentados a esos

nodos, se encuentran m nodos h_i que representan los hospitales. Estos últimos tienen una demanda requerida de p de acuerdo con su instancia basado en sus posiciones disponibles.

Los arcos que unen al nodo 1 con los nodos r_i y los arcos que unen los nodos h_i con el nodo 2 no tienen costo dada su cualidad de transbordo.

La clave del flujo ocurre en los arcos que conectan los nodos r_i con los nodos h_i. Cada residente tendrá m aristas cuyo costo estará definido por la combinación específica residente-hospital pudiendo sumar o restar de acuerdo con la coincidencia con sus listas de preferencias respectivas.

Como cada arco tiene una capacidad de 1 y el nodo de donde se desprende tiene la misma cantidad, la asignación de residente es forzada a optar por el arco más óptimo de acuerdo con la solución general.

Ventajas y desventajas del planteamiento

El planteamiento de este problema presenta la ventaja de ser visual y más adecuado para representar el problema que el de asignación (sobre todo si se presenta a no expertos). No obstante, la mayor desventaja que observamos es la complejidad de realizar un gráfico sencillo ante la magnitud del problema y la definición de las restricciones que resultan menos intuitivas.

Bibliografía consultada: Hillier, F. S. (2007). Introduction to Management Science. Mc Graw-Hill.