Entrega Final Proyecto Optimización

David Santiago Flórez Alsina
davidsa.florez@urosario.edu.co
Estudiante de matemáticas aplicadas y ciencia de la computación.

Alejandra Campo Archbold
alejandra.campo@urosario.edu.co
Economista y Estudiante de matemáticas aplicadas y ciencia de la computación.

Andrey Javier Lizarazo Hernandez
andrey.lizarazo@urosario.edu.co
Estudiante de matemáticas aplicadas y ciencia de la computación.

(Universidad del Rosario)

(Dated: November 24, 2021)

RESUMEN EJECUTIVO

En este proyecto de optimización buscamos obtener la **máxima** cantidad de **ganancia** para una empresa agrícola, teniendo en cuenta el terreno cultivable disponible y buscando no producir un único tipo cultivo. Esto con el fin de **evitar** los problemas que trae el uso de **monocultivos**, y así favorecer la rotación de suelos, y además, **reducir los riesgos** de pérdidas de ingresos de la empresa por la falta de diversificación, problemas climáticos y competencia de mercado [1, 2].

Con esta idea central desarrollamos dos planteamientos del problema, uno para un caso lineal y otro para el caso no lineal. En el primero caso, utilizamos el **algoritmo simplex** para encontrar las cantidades a cultivar de cada producto agrícola de acuerdo a el área de cultivo disponible y un factor de prioridad. Una vez encontradas éstas cantidades por cultivo, hallamos el valor de los beneficios de la empresa con base a los ingresos y costos dados. Esto permite obtener un primer panorama de distribución y optimización de los productos agrícolas, ajustándose a cada problema.

En el segundo caso, consideramos el problema no lineal donde además de los ingresos de producción, hay ingresos por la generación de *hummus* a partir de desechos orgánicos que quedan en el proceso de cultivo. Para solucionar este problema utilizamos una modificación del **método de penalidad** agregando el **método del gradiente** *(también modificado)* para maximizar la función objetivo. En este sentido, observamos que esta robustés del método de descenso del gradiente se *'hereda'* a nuestra adaptación de método de penalidad, dando buenos resultados en términos de la función objetivo y mantentiendo las restricciones que se tenían para el problema.

PLANTEAMIENTO DEL PROBLEMA DE OPTIMIZACIÓN LINEAL

A continuación detallaremos las restricciones y variables del problema no lineal:

- $\vec{X}$: Vector de las cantidades a cultivar en hectáreas, donde cada X_i, corresponde a la cantidad a cultivar de una planta en específico, *(hay n cultivos posibles para plantar).*
- $\vec{I}$: Vector de ingresos por hectárea cultivada de cada cultivo i.
- $\vec{C}$: Vector de costos por hectárea cultivada de tipo de cultivo i.
- b_i: Corresponde a un factor que da prioridad o no a un determinado cultivo. (Este factor toma en consideración el criterio de selección del usuario, es decir, en un caso hipotético si el usuario considera que cierto cultivo no es apropiado para ocupar gran parte del área se le dará un valor bajo, basado en su experiencia, u otros factores que se consideren relevantes).

Así el sistema planteado queda:

$$\begin{aligned}
&\textbf{Max } \vec{I}\cdot\vec{X} - \vec{C}\cdot\vec{X} \\
&\textbf{s.a. } \left(\sum_{i=1}^{n} X_i\right) = \textit{Area total cultivable} \\
&X_i \leq \frac{b_i}{n}\cdot \textit{Area total cultivable} \\
&X_i \geq 0, \quad i = 1, ..., n
\end{aligned}$$

Observe que cada coeficiente b_i es dado por el usuario pero debe cumplir ciertas restricciones básicas que son:

- $b_i \geq 1$, este umbral permite garantizar la participación de todos los cultivos se opcionan como plantables.
- $\sum_{i=1}^{n} b_i > n$, esta restricción sobre la suma de los coeficientes b_i está con el fin de que sea posible llegar a cultivar toda la región cultivable, *(note que esta condición de completar toda el área cultivable también se puede alcanzar con suficiencia si $\sum_{i=1}^{n} b_i = n$ pero en el caso en que todos los b_i son 1 el problema se vuelve trivial).*

RESULTADOS LINEALES

Caso 1

Valores iniciales:

- *Ingresos:* [18 10 5 21], *Costos:* [5 4 2 4], por lo tanto el vector de coeficientes de la función de costo es igual $\boldsymbol{c} = [13\ 6\ 3\ 17]^T$.
- Área total: 20
- Cantidad de cultivos: $n = 4$

Factores de prioridad por cultivo:

- Maíz: $b_1 = 1.2$
- Frijol: $b_2 = 1$
- Papa: $b_3 = 1$
- Albahaca: $b_4 = 1.2$

Luego, el vector de restricciones es igual a $[6\ 5\ 5\ 6\ 20]^T$.

Formato canónico:

$$\begin{array}{ll} \text{Maximizar} & 13x_1 + 6x_2 + 3x_3 + 17x_4 \\ \text{sujeto a} & x_1 \leq 6 \\ & x_2 \leq 5 \\ & x_3 \leq 5 \\ & x_4 \leq 6 \\ & x_1 + x_2 + x_3 + x_4 = 20 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{array}$$

Formato estándar:

$$\begin{array}{ll} \text{Maximizar} & 13x_1 + 6x_2 + 3x_3 + 17x_4 \\ \text{sujeto a} & x_1 + x_5 = 6 \\ & x_2 + x_6 = 5 \\ & x_3 + x_7 = 5 \\ & x_4 + x_8 = 6 \\ & x_1 + x_2 + x_3 + x_4 = 20 \\ & x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8 \geq 0 \end{array}$$

Solución:

Índices obtenidos de la Fase 1:

Índices básicos: [1 2 3 4 5], *Índices no básicos:* [6 7 8].

Resultados de la Fase 2:

El vector solución x es: $[6\quad 5\quad 3\quad 6\quad 0\quad 0\quad 2\quad 0]$

El vector de costos reducidos es: $[3\quad 10\quad 14]$

Los índices básicos finales son: $[1\quad 2\quad 3\quad 4\quad 7]$

CONCLUSIONES CASO 1

En este primer caso, tenemos los productos agrícolas maíz, frijol, papa y albahaca donde maíz y albahaca tiene mayor factor de prioridad respecto a los demás cultivos. Al final los resultados óptimos muestran que maíz, frijol y albahaca deben cultivarse a su mayor restricción posible debido a su relación ingreso-costo. Por otro lado, la papa presenta menores ingresos y mayores costos de tal forma que la cantidad óptima cultivable para este producto es de 3 y no de 5 como se ve reflejado en la restricción.

Finalmente el valor de la función objetivo con el vector solución es 219, es decir, los beneficios de la empresa agricultora son iguales a 219 unidades arbitrarias.

Caso 2

Valores iniciales:

- *Ingresos:* [50 35 28 10], *Costos:* [42 14 15 7], por lo tanto el vector de coeficientes de la función de costo es igual $\boldsymbol{c} = [8\ 21\ 13\ 3]^T$.
- Área total: 20
- Cantidad de cultivos: $n = 4$

Factores de prioridad por cultivo:

- Maíz: $b_1 = 3$
- Frijol: $b_2 = 1.7$
- Papa: $b_3 = 0.8$
- Albahaca: $b_4 = 1.2$

Luego, el vector de restricciones es igual a $[15\ 8.5\ 4\ 6\ 20]^T$.

Formato canónico:

Maximizar	$8x_1 + 21x_2 + 13x_3 + 3x_4$
sujeto a	$x_1 \leq 15$
	$x_2 \leq 8.5$
	$x_3 \leq 4$
	$x_4 \leq 6$
	$x_1 + x_2 + x_3 + x_4 = 20$
	$x_1, x_2, x_3, x_4 \geq 0$

Formato estándar:

Maximizar	$8x_1 + 21x_2 + 13x_3 + 3x_4$
sujeto a	$x_1 + x_5 = 15$
	$x_2 + x_6 = 8.5$
	$x_3 + x_7 = 4$
	$x_4 + x_8 = 6$
	$x_1 + x_2 + x_3 + x_4 = 20$
	$x_i \geq 0,$
	$i = 1, 2, ..., 8$

Solución:

Índices obtenidos de la Fase 1:

Índices básicos: [1 2 3 4 5], *Índices no básicos:* [6 7 8].

Resultados de la Fase 2:

El vector solución x es:
[7.5 8.5 4 0 7.5 0 0 6]

El vector de costos reducidos es: [13 5 5]

Los índices básicos finales son: [1 2 3 5 8]

CONCLUSIONES CASO 2

En este segundo caso, aumentamos los ingresos y los costos de maíz, frijol, papa y disminuimos los ingresos de albahaca. Se mantiene la misma área cultivable disponible. Cambiamos las preferencias del agricultor por su factor de prioridad. Los resultados son acordes al comportamiento de los datos. Las cantidades a cultivar en hectáreas de la albahaca es de cero debido a que sus ingresos son bajos y sus costos aumentaron.

Finalmente, los beneficios de la empresa agricultora son iguales a 295.5 unidades arbitrarias debido al aumento de los ingresos de maíz, frijol, papa.

CONCLUSIONES DEL PROBLEMA LINEAL

El algortimo simplex busca e detectar si ya hemos llegado al óptimo o aún tenemos que pasar a otro punto extremo. También se puede detectar si el problema no tiene solución óptima [3].

El algortimo se ajusta bien a las restricciones y las condiciones del agricultor. El algoritmo se ajusta adecuadamente el área de hectáreas cultivable teniendo en cuenta la prioridad de cada producto agrícola y la relación ingreso-costo. Esto se ve reflejado en las particularidades de cada caso, mejorando la función objetivo.

PLANTEAMIENTO DEL PROBLEMA DE OPTIMIZACIÓN NO LINEAL

A continuación se mostrará las restricciones y variables del problema no lineal:

- $\vec{X}$: En este caso representa la cantidad de árboles x_i del tipo de árbol i a cultivar. En ese sentido, si tenemos árboles de mango por ejemplo, se van a representar como $X_1 =$ "numero de árboles de mango".
- r_i : Representa el radio que abarca un árbol por sus raíces en mts, tomando en cuenta esto, en el radio asignado a un respectivo árbol no se podría plantar otro, ya que ese espacio esta designado para las raíces del árbol en cuestión y para su toma de nutrientes. Por otro lado, entorno al circulo del árbol donde no se puede plantar otro, hay un anillo de maleza el cual se tiene pensado para que ayude a mantener los bichos distraídos y de esa manera no perjudicar el cultivo. Este anillo va a tener un ancho r_i, que depende del tipo de árbol i.
- $H(\vec{X})$: Esto representa una función logarítmica de ingresos. Estos ingresos son de Humus que es materia orgánica degradada a su último estado de descomposición por efectos de microorganismos, que se encuentra químicamente estabilizada, por lo que regula la dinámica de la nutrición vegetal en el suelo (factorhumus, s.f.)[4]. La función tentativamente es de la forma:

$$\vec{v} \cdot \begin{bmatrix} ln(x_1) \\ ln(x_2) \\ ln(x_3) \\ ln(x_4) \end{bmatrix} = \sum_{i=1}^{n} \vec{v}_i \cdot ln(x_i) \tag{1}$$

  donde el vector $\vec{v}$ es $[a\ b\ c\ d]$ donde cada una de estas variables a, b, c, d son constantes que significan que el hummus que proviene de los desechos relacionados al tipo de árbol i es mayor en cantidad. En este sentido $ln(x_i)$ significa la cantidad de ingresos provenientes del hummus dada una cantidad x_i de desechos de un árbol.
- $\vec{I}$: Vector de ingresos por hectárea cultivada de cada árbol i.
- $\vec{C}$: Vector de costos por hectárea cultivada de tipo de árbol i.

$$
\begin{aligned}
&\textbf{Max } (\vec{I}-\vec{C})\cdot\vec{X}+H(\vec{X})\\
&\textbf{s.a. } \sum_{i=1}^{n} x_i\left(r_i+\frac{r_i}{10}\right)^2\pi = \text{ Área total}\\
&x_i \geqslant 1, \quad i=1,\dots,n\\
&x_i \leqslant \frac{b_i\cdot(\text{Área total})}{n\pi\left(r_i+\frac{r_i}{10}\right)^2} \quad i=1,\dots,n
\end{aligned}
$$

IMPLEMENTACIÓN OPTIMIZACIÓN POR PENALIDAD

Para solucionar este problema implementamos una variante del método de penalización ideada por nosotros e inspirada en algunas otras ideas que hemos visto en el curso. En este sentido, nuestro algoritmo modificado funciona a grandes razgos así:

1. Definimos una serie de ecuaciones para cada una de las penalizaciones de desigualdad y de igualdad, y separamos en listas por tipos de penalización *(Los tipos fueron: desigualdades que acotan superiormente, las que acotan inferiormente y las restricciones de igualdad)*, con estas tres categorías logramos realizar un manejo mucho más cómodo de qué castigo impartir y en qué caso.

2. Tras tener estas listas de ecuaciones se las pasamos nuestra adaptación del método del gradiente, que en resumen lo que hace es revisar para el punto actual qué restricciones se incumplen, con base a esto selecciona su respectiva ecuación, le aumenta la penalidad ρ_i a esta ecuación en particular y la añade a la función objetivo *(hay un factor de penalidad particular para cada restricción)*.

3. Tras actualizar los parámetros cuando se añade la nueva penalización, se vuelve a correr el algoritmo y se repite el proceso.

4. Para encontrar el mejor óptimo *(o al menos estar más cerca)* decidimos correr con 40 puntos iniciales distintos el algoritmo, y para evitar esperar demasiado tiempo decidimos paralelizar la ejecución del algoritmo obteniendo grandes resultados *obteníamos los resultados hasta 17 veces más rápido en comparacion con la versión serial.*

5. Tras obtener las 40 respuestas para los 40 distintos puntos iniciales buscabamos cuál era el vector solución que mejores resultados nos daba en términos de la función objetivo, y tomábamos este por solución.

Los pasos 3 y 4 los implementamos puesto que no podemos asumir mucho sobre la topología de la región sobre la cual vamos a optimizar, y al tener varios puntos iniciales garantiza un plus de resiliencia ante posibles mínimos locales.

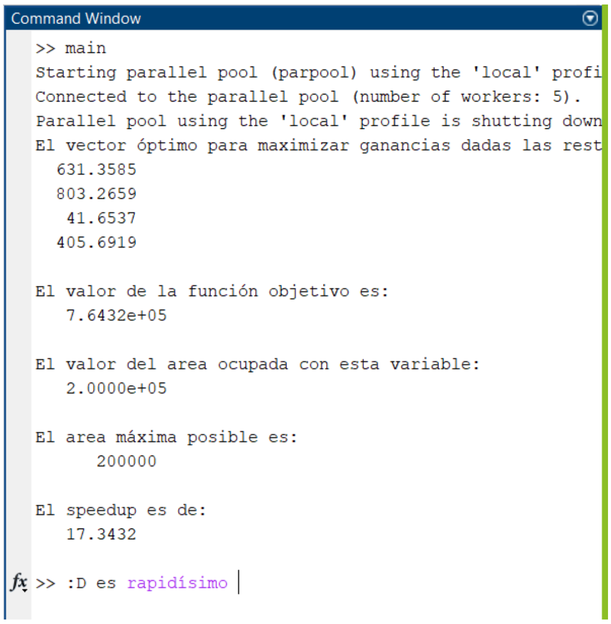

FIG. 1: *Resultados obtenidos en un procesador Ryzen 5 3600X, con 8 GB de ram de 3200 MHZ.*

RESULTADOS PROBLEMA NO LINEAL

En este caso simulamos la plantación de *cerezos, aguacate, mango y durazno.* Para probar nuestro algoritmo tomamos varios escenarios, sin embargo hubo una serie de parámetros que se mantubieron constantes en entre estas variaciones, estos fueron:

$$
\begin{aligned}
r &= [5 \;\; 4 \;\; 15 \;\; 6]\\
\vec{v} &= [2 \;\; 1 \;\; 3 \;\; 2]
\end{aligned}
$$

Las unidades de radio de los árboles están en metros, y el area total sembrable está en metros cuadrados, por lo que pasamos de 20 ha a 200.000 mts^2. Los casos que asumimos fueron variaciones las ganancias lineales de los cultivos, con la intención de ver como se comportaba el algoritmo ante cambios en la función objetivo:

Caso 1

Los valores de los ingresos, los costos y los factores de prioridad por cultivo fueron:

$$
\begin{aligned}
\vec{I} &= \begin{bmatrix}400 & 500 & 200 & 400\end{bmatrix}\\
\vec{C} &= \begin{bmatrix}20 & 50 & 10 & 18\end{bmatrix}
\end{aligned}
$$

El vector de factores de prioridad es:

$$
b = \begin{bmatrix}1.2 & 1 & 1.3 & 1.3\end{bmatrix}
$$

Con estos coeficientes las cantidades máximas de árboles que pueden plantarse de cada tipo son:

$$\begin{bmatrix} 631.3585 \\ 822.0813 \\ 75.9969 \\ 474.9803 \end{bmatrix}$$

El vector solución fue:

$$x^* = \begin{bmatrix} 515.7247 \\ 807.7704 \\ 43.0954 \\ 474.9803 \end{bmatrix}$$

El valor de la función objetivo es: 7.4915e+05
El valor del area ocupada con esta variable: 2.0000e+05

Caso 2

Los valores de los ingresos, los costos y los factores de prioridad por cultivo fueron:

$$\vec{I} = \begin{bmatrix} 300 & 600 & 200 & 400 \end{bmatrix}$$
$$\vec{C} = \begin{bmatrix} 100 & 750 & 100 & 18 \end{bmatrix}$$

El vector de factores de prioridad es:

$$b = \begin{bmatrix} 1.2 & 1 & 1.3 & 1.3 \end{bmatrix}$$

Con estos coeficientes las cantidades máximas de árboles que pueden plantarse de cada tipo son:

$$\begin{bmatrix} 631.3585 \\ 822.0813 \\ 75.9969 \\ 474.9803 \end{bmatrix}$$

La solución fue:

$$x^* = \begin{bmatrix} 631.3585 \\ 392.0017 \\ 75.9969 \\ 353.0691 \end{bmatrix}$$

El valor de la función objetivo es: 2.0999e+05 El valor del area ocupada con esta variable: 1.9716e+05

CONCLUSIONES DEL PROBLEMA NO LINEAL

A lo largo de las distintas iteraciones del código pudimos observar muchos cambios y entender mucho mejor el algoritmo, históricamente comenzamos por mantener un factor de penalización ρ constante para todas las restricciones, lo que causaba que el algoritmo no se adaptara tanto a los cambios en la función objetivo que construimos, por otra parte, en las versiones iniciales aplicábamos la penalización de restricciones de desigualdad en todo momento, es decir no distinguíamos entre si se estaban cumpliendo las restricciones de desigualdad *(y por lo tanto no debían castigarse)* y el caso en que se incumplían *(había que castigar las particulares que no eran cumplidas)*.

Lo anterior causó un fenómeno muy interesante y es que dado que en todo momento estaban presentes las restricciones de desigualdad, el algoritmo trataba de no alejarse demasiado de ninguna de estas restricciones porque la penalización para alejarse sería cuadrática y multiplicada por un ρ bastante alto que habíamos dejado al inicio. Esto en resumen nos causó una falta de *'plasticidad'* en la función objetivo, ya que en todo momento el pobre algoritmo se enfocaba en no ser castigado por las restricciones de desigualdad.

Con esto entendido y algunos comentarios que recibimos, nos pusimos a la tarea de modificar el algoritmo para que tuviese factores ρ de penalización adaptables y que además, tuviese en cuenta ese *'contexto'* de qué variables estaban cumpliendo la penalización y no debían ser castigadas dentro de la función objetivo.

Una de las cosas interesantes que notamos y que al mismo tiempo eran algo esperadas es que cuando corremos el algoritmo con mismos datos, el algoritmo nos retorna distintas respuestas, cabe destacar que no son demasiado distintas pero es importante resaltar esta caracterísitica que surge de generar puntos aleatorios para correr el algoritmo inicialmente.

Por otra parte para los casos que probamos pudimos verificar que se cumplen todas las restricciones, tanto las de desigualdad como las de igualdad, esta última es crítica ya que no podemos bajo ningún motivo plantar más árboles de los que permite el espacio.

[1] I. A. and T. A., Linear Programming in Agriculture: Case Study in Region of Development South-Mountenia. (English), International Journal of Substainable Economies Management. , 51 (2012).
[2] M. S. Scarpari and E. G. F. de Beauclair, Optimized agricultural planning of sugarcane using linear programming, Investigacion operacional **31**, 126 (2010).
[3] J. R. Berrendero, Optimización lineal. algoritmo del simplex, .
[4] F. Humus, Revolucionario humus de lombriz, , `https://www.factorhumus.com/` (s.f.).