El Método de Conjunto Activo para Problemas de Optimización Cuadráticos

Adrian Tame Jacobo
Estudiate de Matemáticas Aplicadas del ITAM

Introducción

El mundo de la optimización empieza mucho antes que el nacimiento de las computadoras, con su concepto moderno introducido por Alan Turing en su ya famoso artículo [1]. Fermat y Lagrange son generalmente acreditados por presentar los primeros resultados importantes, utilizando resultados de Cálculo para poder identificar puntos óptimos de un problema, mientras que Newton y Gauss crearon métodos iterativos para poder encontrar una solución a problema de optimización. El método de Newton, una base central de muchos algoritmos de optimización, es presentado a detalle en [2].
Hoy en día, los problemas de optimización que se intentan resolver son mucho más complejos que hace tiempo, y se necesitan nuevas herramientas para poder resolverlos. Una de las formas interesantes que aparecen en problemas de optimización es el problema cuadrático, definido formalmente como

$$\begin{aligned} \underset{x}{\text{minimizar}} \quad & q(x) = \frac{1}{2}x^T Q x + x^T c \\ \text{sujeto a} \quad & a_i^T x = b_i,\ i \in \mathcal{E} \\ & a_i^T x \geq b_i,\ i \in \mathcal{I} \end{aligned} \tag{1}$$

con Q una matriz simétrica de $n \times n$, $\mathcal{I}, \mathcal{E}$ conjuntos finitos de índices, y $x, c, \{a_i\}, i \in \mathcal{I} \cup \mathcal{E}$ vectores en $\mathbb{R}^n$ [3]. Es de notar que las restricciones de este problema son puramente lineales. En la siguiente sección, se explora este problema y sus varias aplicaciones.

Programación Cuadrática [1]

El problema (1) es uno muy conocido y estudiado en el área de investigación de operaciones. Es también un subproblema importante en los algoritmos de programación cuadrática secuencial, métodos de Lagrange aumentados, y métodos de punto interior. En general, muchas de las aplicaciones de este problema caen en las áreas de finanzas, agricultura, producción y operaciones, economía y marketing.
Probablemente una de las razones por las cuales este problema es tan conocido es porque es una de las formas más simples de problemas no lineales. Los problemas de esta forma siempre pueden ser resueltos en una cantidad finita de iteraciones de un método, o en el caso contrario se puede demostrar que no tienen soluciones factibles. Se nota que el método de conjunto activo que se expone aquí no garantiza convergencia al problema general, pero sí a ciertos subcasos del problema, como por ejemplo el caso en el que solamente se

[1] La mayoría de esta información viene de [3] y [4].

tienen restricciones de desigualdad.
Por completitud, y porque es un tema bastante interesante, continúan algunos casos detallados donde el problema se modela como un problema de la forma de (1).

- Markowitz [5] desarrolló aplicaciones importantes en finanzas basadas en modelos cuadráticos. Estos son para modelar portafolios de inversión sujetos a algún riesgo.

- Hay algunas aplicaciones interesantes a optimización de formas geométricas para optimizar el espacio utilizado al cortar diamantes, y gastar la menor cantidad posible de materia prima al cortar.

- El modelo de cambio de marca basado en procesos de Markov utiliza estos métodos. Theil y Ray desarrollaron un modelo para el cambio de un cliente a otro y se estima con un modelo cuadrático.

- Louwes, Boot y Wage desarrollaron un modelo para el uso óptimo de la leche en Países Bajos. Este modelo es usado en muchas aplicaciones de agricultura.

- Otra aplicación interesante es la de separación lineal. Se puede modelar un problema de separación lineal como la distancia en norma de un vector a un hiper-plano lineal, y minimizar las distancias para generar un hiper-plano que separe mejor los datos. La siguiente imagen es una implementación de esto al problema de Iris [6] con solamente 3 variables consideradas, que son la longitud del sépalo, ancho del sépalo y ancho del pétalo.

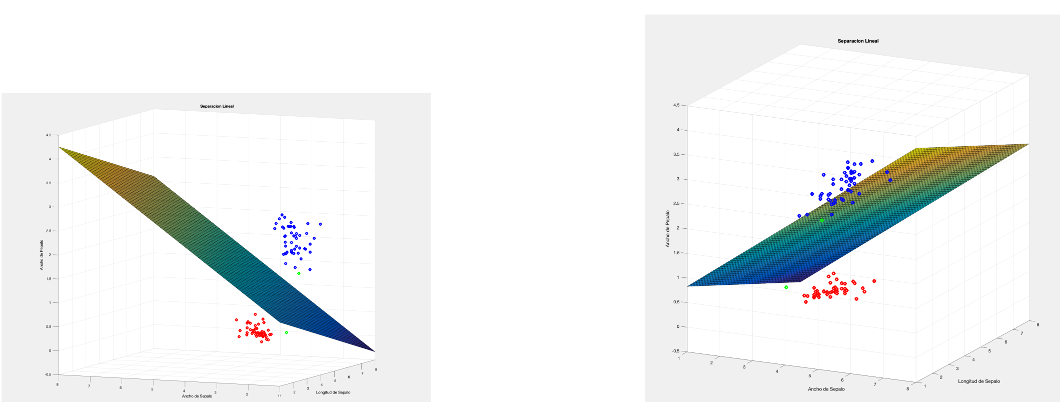

Figura 1: Separación lineal del conjunto de datos Iris.

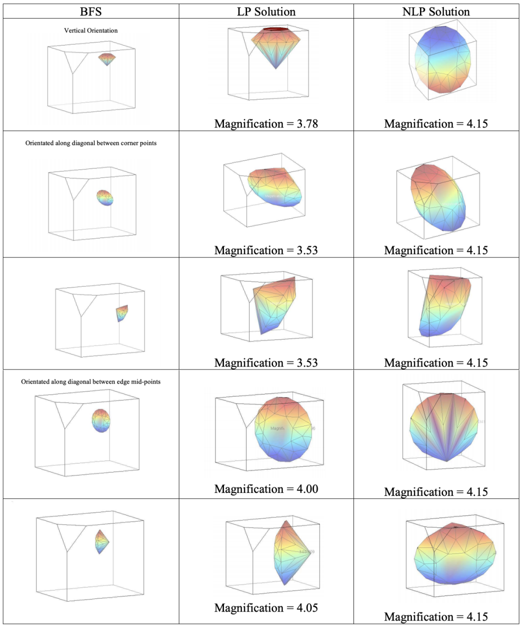

Figura 2: Corte de diamantes óptimo, tomado de [7]

El Método de Conjunto Activo

En principio, lo que hace este método es tomar un punto x, y checar qué condiciones de desigualdad se cumplen en ese punto de manera estricta. Esto es algo interesante, ya que el algoritmo eventualmente (en el caso de convergencia) convergerá a un punto, el cual nos puede decir mucho de la naturaleza del problema a resolver. Por ejemplo, saber cuáles restricciones están activas nos permite saber cuáles de ellas son las de mayor importancia o juegan un papel más importante en la minimización del problema.
El método resuelve un subproblema (este se puede ver de manera clara en el pseudocódigo, dado por (2)). En general, se trabaja en un conjunto convexo ya que las restricciones están dadas por igualdades y desigualdades. También, una condición importante para poder resolver el problema es que Q sea una matriz positiva definida y simétrica, pero en algunos casos con matrices indefinidas también se puede resolver, sin embargo, esto está fuera del alcance de este artículo.

Pseudocódigo

A continuación se presenta el pseudocódigo para resolver el problema (1) utilizando el método de conjunto activo. Este algoritmo está fuertemente basado en el algoritmo de conjunto activo presentado en [3]. Además, esta implementación considera $\mathcal{E} = \emptyset$.

Result: **x** un vector solución de tamaño n y μ un vector con los multiplicadores de Lagrange asociados.
Inicialización: Sea Q una matriz de tamaño n x n simétrica, positiva definida;
Sea c un vector de tamaño n;
Sea μ un vector de tamaño n el vector de multiplicadores de lagrange completo ;
Sea λ un vector de tamaño cero inicialmente ($\emptyset$);
Sea F una matriz de coeficientes de restricciones de tamaño m x n con rango(F) = $m \leq n$;
Sea d un vector de valores de restricciones de tamaño m;
while iteraciones no excedan un limite arbitrario **do**
  Sea $g_k = Qx_k + c$;
  Sea $\mathcal{W}_k$ el conjunto de condiciones activas en el punto x_k;
  resolver el subproblema dado por:

$$\begin{aligned} \underset{p}{\text{minimizar}} \quad & \frac{1}{2} p_k^T Q p_k + g_k^T p_k \\ \text{sujeto a} \quad & a_i^T p_k = 0, \ i \in \mathcal{W}_k \end{aligned} \tag{2}$$

  if $||p_k||_2 = 0$ **then**
    if $\lambda \leq 0$ **then**
      END;
    else
      $j = argmin\{j \in I \cap \mathcal{W}_k, \hat{\mu_j}\}$;
      Actualizar x_k a x_{k+1};
      $\mathcal{W}_{k+1} = W_k - \{j\}$
    end
  else
    Activos = Todas las condiciones que cumplan que $Fx_k = d$;
    $\mathcal{W}_c$ = Activos $-\mathcal{W}_k$;
    if Hay una condición que cumpla que $F(\mathcal{W}_k)p = 0$ **then**
      $\mathcal{V} = \mathcal{W}_c[F(\mathcal{W}_c)] \cdot p_k = 0)$;
      $j = min\,\{F(\mathcal{W}_c) \cdot p_k = 0\}$;
      $\mathcal{W}_k = \mathcal{W}_k \cup \mathcal{V}_j$;
    else
      Hacer búsqueda de linea con condiciones de Wolfe para encontrar $_k$;
      $x_k + 1 = x_k + \alpha_k p_k$;
    end
  end
  $\mu(Activas) = -\lambda$;
  $\mu(\mathcal{W}_c) = 0$;
end

Referencias

[1] A. M. Turing; On Computable Numbers, with an Application to the Entscheidungsproblem, *Proceedings of the London Mathematical Society*, Volumen s2-42, Número 1, 1° de Enero de 1937, Páginas 230–265, https://doi.org/10.1112/plms/s2-42.1.230

[2] Polyak, Boris. (2007). Newton's method and its use in optimization. *European Journal of Operational Research*. 181. 1086-1096. https://doi.org/10.1016/j.ejor.2005.06.076.

[3] Nocedal, J. and Wright, S.J. (1999) Numerical Optimization. Springer, New York. https://doi.org/10.1007/b98874

[4] Omprakash K. Gupta (1995) Applications of Quadratic Programming, Journal of Information and Optimization Sciences, 16:1, 177-194, https://doi.org/10.1080/02522667.1995.10699213

[5] H. M. Markowitz (1952), Portfolio selection, Journal of Finance, Vol. 7, pp. 77-91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

[6] http://archive.ics.uci.edu/ml/datasets/iris

[7] Viswambharan, Ambili. (1998). Optimisation in Diamond Cutting.