IN77O - Modelos y Algoritmos de Optimización

Resumen Métodos de punto interior

Basado en el Capítulo 11 del libro *Convex Optimization*

Profesor: Fernando Ordóñez
Auxiliares: Benjamín Barrientos, José Miguel González

11. Métodos de punto interior

11.1. Minimización con restricciones de desigualdad

En lo que sigue, se discuten métodos para resolver problemas de optimización que incluyen restricciones de igualdad:

$$\begin{array}{ll} \min & f_0(\mathbf{x}) \\ \text{s.a.} & f_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m \\ & \mathbf{Ax} = \mathbf{b} \end{array}$$

con $f_0, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$ convexas y dos veces continuamente diferenciables, $\mathbf{A} \in \mathbb{R}^{p \times n}$ con rango$(\mathbf{A}) = p < n$ y $\mathbf{b} \in \mathbb{R}^p$. Asumiremos que existe un punto óptimo $\mathbf{x}^*$, y denotaremos el valor óptimo $f(\mathbf{x}^*)$ como p^*. También asumiremos que se cumple la condición de Slater, por tanto existen óptimos duales $\boldsymbol{\lambda}^* \in \mathbb{R}^m$, $\boldsymbol{\nu}^* \in \mathbb{R}^p$ que junto a $\mathbf{x}^*$ satisfacen las condiciones de KKT:

$$\mathbf{Ax}^* = \mathbf{b}, \qquad f_i(\mathbf{x}^*) \leq 0,\ i = 1, \ldots, m$$

$$\boldsymbol{\lambda}^* \succeq \mathbf{0}$$

$$\nabla f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \boldsymbol{\lambda}_i^* \nabla f_i(\mathbf{x}^*) + \mathbf{A}'\boldsymbol{\nu}^* = \mathbf{0}$$

$$\boldsymbol{\lambda}_i^* f_i(\mathbf{x}^*) = 0,\ i = 1, \ldots, m$$

Los métodos de punto interior resuelven el problema aplicando el Método de Newton a una secuencia de problemas con restricciones de igualdad. En este resumen nos concentraremos en un algoritmo de punto interior particular, el método de barrera.

11.2. Función de barrera logarítmica y camino central

El objetivo ahora es formular aproximadamente el problema inicialmente presentado como un problema con restricciones de igualdad al cual podamos aplicar el Método de Newton. El primer paso es escribir las restricciones de desigualdad en forma implícita en la función objetivo:

$$\begin{array}{ll} \min & f_0(\mathbf{x}) + \sum_{i=1}^{m} I_-(f_i(\mathbf{x})) \\ \text{s.a.} & \mathbf{Ax} = \mathbf{b} \end{array}$$

donde $I_- : \mathbb{R} \to \mathbb{R}$ es una función indicatriz para los reales no positivos:

$$I_-(u) = \begin{cases} 0 & \text{si} \quad u \leq 0 \\ \infty & \text{si} \quad u > 0 \end{cases}$$

Notamos que este nuevo problema no tiene restricciones de desigualdad, pero su función objetivo en general no es diferenciable.

11.2.1. Barrera logarítmica

La idea central del método de barrera es aproximar la función I_- por:

$$\hat{I}_-(u) = -\frac{1}{t}\log(-u), \qquad \text{dom } \hat{I}_- = -\mathbb{R}_{++}$$

donde t es un parámetro que define la precisión de la aproximación: a mayor t, mejor aproximación. Al igual que I_-, $\hat{I}_-$ es convexa y no decreciente, pero a diferencia de ella es diferenciable. El problema aproximado con $\hat{I}_-$ queda:

$$\begin{array}{rl} \min & f_0(\mathbf{x}) + \sum_{i=1}^{m} -\frac{1}{t}\log(-f_i(\mathbf{x})) \\ \text{s.a.} & \mathbf{Ax} = \mathbf{b} \end{array}$$

La función

$$\phi(\mathbf{x}) = -\sum_{i=1}^{m}\log(-f_i(\mathbf{x})),$$

con dominio dom $\phi = \{\mathbf{x} \in \mathbb{R}^n \mid f_i(\mathbf{x}) < 0,\ i = 1, \dots, m\}$, se llama la **barrera logarítmica** del problema.

Es importante notar que la elección de t genera un *trade-off*: mientras mayor es, mejor es la calidad de la aproximación (veremos esto más adelante), pero más difícil es resolver el problema usando método de Newton, porque el Hessiano de la función objetivo varía muy bruscamente cerca de la frontera de la región factible. Veremos que esto se puede sobrellevar resolviendo una secuencia de problemas, aumentando el parámetro t en cada paso y comenzando cada aplicación del método de Newton en la solución del problema para el valor previo de t.

11.2.2. Camino central

En lugar del problema aproximado ya descrito, en general consideraremos el siguiente problema equivalente que se obtiene al multiplicar la función objetivo por t:

$$\begin{array}{rl} \min & tf_0(\mathbf{x}) + \phi(\mathbf{x}) \\ \text{s.a.} & \mathbf{Ax} = \mathbf{b} \end{array}$$

Asumiremos que este problema puede ser resuelto mediante el método de Newton y que tiene una única solución $\mathbf{x}^*(t)$ para cada $t > 0$.

Definición 11.1. El **camino central** asociado al problema es el conjunto $\{\mathbf{x}^*(t) \mid t > 0\}$.

Los puntos del camino central están caracterizados por las siguientes condiciones necesarias y suficientes:

- Factibilidad primal:
$$\mathbf{Ax}^*(t) = b, \qquad f_i(\mathbf{x}^*(t)) < 0,\ i = 1, \dots, m$$
- Gradientes alineados: existe $\hat{\boldsymbol{\nu}} \in \mathbb{R}^p$ tal que
$$0 = t\nabla f_0(\mathbf{x}^*(t)) + \nabla\phi(\mathbf{x}^*(t)) + \mathbf{A'}\hat{\boldsymbol{\nu}} = t\nabla f_0(\mathbf{x}^*(t)) + \sum_{i=1}^{m}\frac{1}{-f_i(\mathbf{x}^*(t))}\nabla f_i(\mathbf{x}^*(t)) + \mathbf{A'}\hat{\boldsymbol{\nu}}$$

Teorema 11.1. Todo punto $\mathbf{x}^*(t)$ perteneciente al camino central verifica

$$f_0(\mathbf{x}^*(t)) - p^* \leq \frac{m}{t},$$

es decir, es a lo más m/t-subóptimo. Esto confirma la idea de que $\mathbf{x}^*(t) \xrightarrow{t\to\infty} \mathbf{x}^*$.

11.3. El método de barrera

Del resultado de la sección anterior, podría pensarse en definir una tolerancia ε, fijar $t = m/\varepsilon$ y resolver el problema aproximado para este valor del parámetro t, obteniendo una solución ε-subóptima. Sin embargo, para problemas grandes este método no funciona en forma adecuada, por lo cual prácticamente no se utiliza.

11.3.1. El método de barrera

Una extensión muy sencilla del método anterior, que consiste en resolver el problema aproximado para valores crecientes del parámetro t, comenzando el método de Newton en cada iteración con el óptimo del problema anterior y parando al llegar a $t \geq m/\varepsilon$, sí funciona bien. Este algoritmo se describe a continuación.

Algorithm 1 Método de punto interior

given Punto de partida estrictamente factible $\mathbf{x}$, $t = t^{(0)} > 0$, parámetro de crecimiento de t $\mu > 1$ y parámetro de tolerancia ε
while true do
  Calcular $\mathbf{x}^*(t)$ con método de Newton, minimizando $tf_0 + \phi$ sujeto a $\mathbf{A}\mathbf{x} = \mathbf{b}$, comenzando en $\mathbf{x}$
  $\mathbf{x} := \mathbf{x}^*(t)$
  if $(\frac{m}{t} < \varepsilon)$ **then**
    break
  end if
  $t := \mu t$
end while

La elección de μ y de $t^{(0)}$ conllevan *trade-offs*. Por un lado, si μ es muy cercano a 1 el método de Newton resulta más rápido porque se cuenta con un mejor punto inicial, pero el número de iteraciones tiende a ser más grande porque el *gap* de optimalidad se reduce muy poco en cada una de ellas. Para μ demasiado grande ocurre lo contrario: el método de Newton tarda más, pero se requieren menos iteraciones para alcanzar el valor de t deseado. En general se utilizan valores entre 10 y 20. En cuanto a $t^{(0)}$, el *trade-off* es directo: si es muy grande, encontrar el primer punto del camino central será muy complejo; si es muy pequeño, se requerirán más iteraciones del método.

11.3.3. Análisis de convergencia

Teorema 11.2. El método de barrera para un problema con m restricciones de desigualdad, con parámetro inicial $t^{(0)}$, parámetro de aumento μ y parámetro de tolerancia ε, finaliza después de exactamente

$$\left\lceil \frac{\log(m/(\varepsilon t^{(0)}))}{\log \mu} \right\rceil$$

iteraciones, además de la iteración inicial.

Referencias

[1] Stephen Boyd y Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, mar. de 2004. ISBN: 0521833787. URL: http://www.amazon.com/exec/obidos/redirect?tag=citeulike-20%5C&path=ASIN/0521833787.