Método del Lagrangiano aumentado

Sharis Barrios García - Manuel Martínez Flores

Noviembre 2024

Contenido

Problemas Generales de Optimización No Lineal

Los problemas generales de optimización no lineal son de la forma

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{Sujeto a} & h_i(x) = 0, \quad i = 1, \ldots, l, \\ & g_j(x) \leq b_j, \quad j = 1, \ldots, p \end{array}$$

Problemas Generales de Optimización No Lineal

Los problemas generales de optimización no lineal son de la forma

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{Sujeto a} & h_i(x) = 0, \quad i = 1, \ldots, l, \\ & g_j(x) \leq b_j, \quad j = 1, \ldots, p \end{array}$$

donde las funciones $f : \mathbb{R}^k \to \mathbb{R}$, $h : \mathbb{R}^k \to \mathbb{R}^l$, $g : \mathbb{R}^k \to \mathbb{R}^p$ son dos veces continuamente diferenciables en $\mathbb{R}^n$.

Problemas Generales de Optimización No Lineal

Notemos que se puede reescribir como:

$$c_i(x) = \begin{cases} h_i(x) & i = 1, ..., l \\ \max(g_{i-l}(x) - b_{i-l}, 0) & i = l+1, ..., l+p \end{cases}$$

Por lo que se puede tomar el problema anterior como:

$$\begin{aligned} \text{Minimizar} \quad & f(x) \\ \text{Sujeto a} \quad & c(x) = 0 \end{aligned}$$

Problemas Generales de Optimización No Lineal

La función ϕ se denomina función objetivo y el conjunto:

$$\Omega = \{x \in \mathbb{R}^n : c(x) = 0\}$$

se llama conjunto factible.

Problemas Generales de Optimización No Lineal

La función ϕ se denomina función objetivo y el conjunto:

$$\Omega = \{x \in \mathbb{R}^n : c(x) = 0\}$$

se llama conjunto factible. Y una solución del problema es un punto $x^* \in \Omega$ tal que $f(x^*) \leq f(x)$ para todo $x \in \Omega$.

Solución

En 1969, Powell y Hestenes propusieron el método del Lagrangiano aumentado como una alternativa más robusta en el contexto de problemas de programación no lineal, cuando se buscaban métodos que permitieran manejar restricciones de una forma más efectiva y estable que los métodos de penalización tradicionales.

Métodos de Penalización Cuadrática

Uno de los métodos de penalización tradicionales ampliamente usados es el método de penalización cuadrática. El cual, agrega términos de penalización al objetivo para forzar que las restricciones se cumplan/ castigar las violaciones de las restricciones. Pero pueden presentar problemas de estabilidad y precisión.

Métodos de Penalización Cuadrática

Def.

Método de penalización cuadrática con restricciones de igualdad:

$$\min_{x \in \mathbb{R}^n} \phi(x, \rho) = f(x) + \frac{\rho}{2}\|c(x)\|^2$$

Métodos de Penalización Cuadrática

Def.

Método de penalización cuadrática con restricciones de igualdad:

$$\min_{x \in \mathbb{R}^n} \phi(x, \rho) = f(x) + \frac{\rho}{2} \|c(x)\|^2$$

donde $\rho > 0$ es un parámetro de penalización. El objetivo es hacer que los valores de $c_i(x)$ se acerquen a cero, penalizando más fuertemente a medida que ρ crece. Por lo que generalmente se elige una suseción $\rho_k \to \infty$ y se utiliza la solución de $\phi(x, \rho_k)$ como posición inicial para solucionar $\phi(x, \rho_{k+1})$

Algoritmo Penalización de cuadrados

Algorithm 1 Penalización de cuadrados

for $k = 1, 2, \ldots$ **do**
  if convergencia **then**
    parar
  end if
  $x_k \leftarrow \text{argmin}_x \phi(x, \rho_k)$ con valor inicial x_{k-1}
  escoger $\rho_{k+1} > \rho_k$
end for

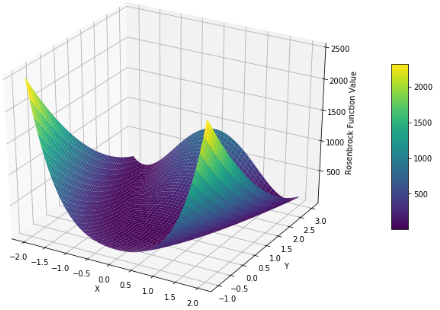
Convergencia de Penalización Cuadrática

Teorema

Si x_k es el mínimo global de $\phi(x, \rho_k)$ con $\rho_k \uparrow \infty$, entonces cualquier punto límite x^* de esa secuencia es mínimo global factible de f

Ejemplo

Considérese la función de Rosenbrock $f(x,y) = (1-x)^2 + 100(y - x^2)^2$ sin restricciones.



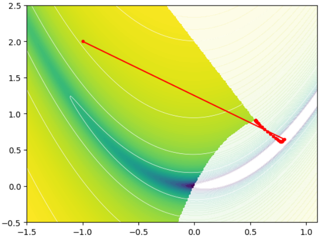
Entonces, la función tiene un mínimo global en $(x,y) = (1,1)$, donde $f(x,y) = 0$. En el caso trivial de $a = 0$ la función es simétrica y el mínimo está en el origen.

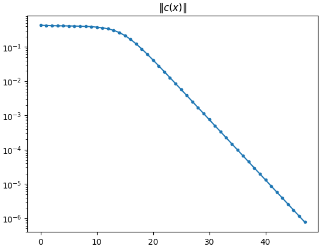
Ejemplo

Considérese la función de Rosenbrock $f(x,y) = (1-x)^2 + 100(y - x^2)^2$ con $(x-1)^3 - y + 1 < 0$ y $2x - y - 2 < 0$. Con valor inicial $(x,y) = (-1,2)$ y $\rho_0 = 1$

Ejemplo

Considérese la función de Rosenbrock $f(x, y) = (1-x)^2 + 100(y - x^2)^2$ con $(x-1)^3 - y + 1 < 0$ y $2x - y - 2 < 0$. Con valor inicial $(x, y) = (-1, 2)$ y $\rho_0 = 1$





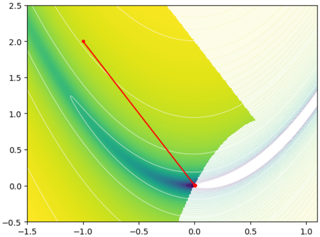
Nótese que se necesita un $\rho \approx 283{,}387{,}333$ para obtener una factibilidad aceptable $\|c(x, y)\| <$ 1e-6

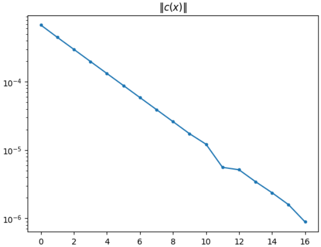
Ejemplo

Considérese la función de Rosenbrock $f(x,y) = (1-x)^2 + 100(y - x^2)^2$ con $(x-1)^3 - y + 1 < 0$ y $2x - y - 2 < 0$. Con valor inicial $(x,y) = (-1,2)$ y $\rho_0 = 1000$

Ejemplo

Considérese la función de Rosenbrock $f(x, y) = (1 - x)^2 + 100(y - x^2)^2$ con $(x - 1)^3 - y + 1 < 0$ y $2x - y - 2 < 0$. Con valor inicial $(x, y) = (-1, 2)$ y $\rho_0 = 1000$





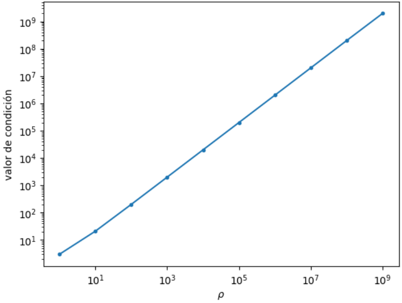
Nótese que se necesita un $\rho \approx 985,261$ para obtener una factibilidad aceptable $\|c(x, y)\| <$ 1e-6

Ejemplo

Considérese $f(x, y) = -\ln x - \ln y$ con $x + y < 1$.
Utilizando valor inicial $(x, y) = (0.3, 0.3)$

Ejemplo

Considérese $f(x, y) = -\ln x - \ln y$ con $x + y < 1$.
Utilizando valor inicial $(x, y) = (0.3, 0.3)$
En este caso, no converge el algoritmo :(Se puede observar que el número de condición de la Hessiana es proporcional al parámetro ρ.



Lagrangiano Estándar y Multiplicadores de Lagrange

Dado un problema de la forma:

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{Sujeto a} & h_i(x) = 0, \quad i = 1, \ldots, l, \\ & g_j(x) \leq b_j, \quad j = 1, \ldots, p \end{array}$$

se reescriben las condiciones como:

$$c_i(x) = \begin{cases} h_i(x) & i = 1, ..., l \\ \max(g_{i-l}(x) - b_{i-l}, 0) & i = l+1, ..., l+p \end{cases}$$

Y el Lagrangiano estandar es:

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T c(x),$$

donde λ_i son los multiplicadores de Lagrange. La función $\mathcal{L}(x, \lambda)$ permite transformar el problema con restricciones en un problema sin restricciones.

Lagrangiano Aumentado

El Langraniano Aumentado es un método que combina ideas de los multiplicadores de Lagrange (usados en optimización con restricciones de igualdad) con términos de penalización cuadrática explicados anteriormente, logrando cumplir mejor las restricciones y proporcionando una convergencia más estable.

Comparación: Penalización Cuadrática vs. Lagrangiano Aumentado

El método de penalización cuadrática tiende a ser inestable al aumentar el parámetro ρ, ya que grandes valores de penalización pueden afectar la condición numérica del problema. En cambio, el Lagrangiano aumentado combina los términos de penalización con los multiplicadores de Lagrange, logrando una convergencia más estable al no depender de un valor elevado de ρ.

Comparación entre Lagrangiano Estándar y Lagrangiano Aumentado

El Lagrangiano estándar es útil para manejar restricciones de igualdad, pero no controla bien las violaciones de las restricciones, ya que no penaliza las desviaciones de $c(x)$ fuera de cero. El Lagrangiano aumentado, por otro lado, incluye un término de penalización cuadrática:

$$\mathcal{L}_A(x, \lambda, \rho) = f(x) + \lambda^T c(x) + \frac{\rho}{2}\|c(x)\|^2,$$

donde $\rho > 0$ es un parámetro de penalización. Este término penaliza las violaciones de las restricciones y mejora la convergencia hacia el conjunto factible.

Lagrangiano Aumentado

Entonces, el Lagrangiano Aumentado busca minimizar la función

$$\text{Minimizar } \mathcal{L}_A(x, \lambda, p) \quad \text{sujeto a } x \in \Omega$$

Donde,

$$c_i(x) = \begin{cases} h_i(x) & i = 1, ..., l \\ \max(g_{i-l}(x) - b_{i-l}, 0) & i = l+1, ..., l+p \end{cases}$$

De esta forma, se observa que este método tiene como objetivo incorporar las restricciones dentro de la función objetivo.

Algoritmo Lagrangiano Aumentado

Algorithm 2 Lagrangiano Aumentado

```
for k = 1, 2, ... do
    if convergencia then
        parar
    end if
    x_k ← argmin_x L_A(x, λ_k, ρ_k) con valor inicial x_{k-1}
    λ_{k+1} ← λ_k + ρ_k c(x)
    escoger ρ_{k+1} ≥ ρ_k
end for
```

Convergencia Lagrangiano Aumentado

Teorema

Sea x^* solución local del problema en el que las derivadas de las restricciones son l.i., y que el criterio de segunda derivada concuerda con $\lambda*$. Entonces existe $\overline{\rho}$ tal que si $\rho \geq \overline{\rho}$, entonces x^* es mínimo local estricto de $\mathcal{L}_A(x, \lambda^*, \rho)$
Además, existe un escalar M tal que

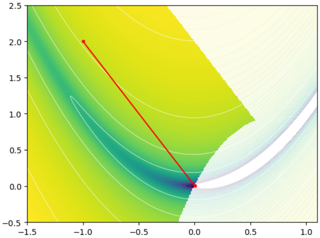
- $\mathcal{L}_A(x, \lambda_k, \rho_k)$ tiene solución factible única x_k y $\|x_k - x^*\| \leq M\|\lambda_k - \lambda^*\|/\rho_k$
- Si $\rho_k \geq \overline{\rho}$, entonces $\|\lambda_{k+1} - \lambda^*\| \leq M\|\lambda_k - \lambda^*\|/\rho_k$

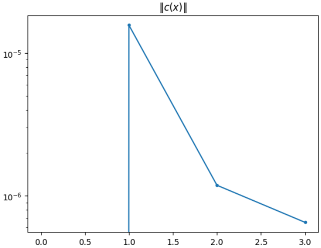
Ejemplo

Considérese la función de Rosenbrock $f(x,y) = (1-x)^2 + 100(y-x^2)^2$ con $(x-1)^3 - y + 1 < 0$ y $2x - y - 2 < 0$. Con valor inicial $(x,y) = (-1,2)$ con $\rho_0 = 1000$

Ejemplo

Considérese la función de Rosenbrock $f(x, y) = (1 - x)^2 + 100(y - x^2)^2$ con $(x - 1)^3 - y + 1 < 0$ y $2x - y - 2 < 0$. Con valor inicial $(x, y) = (-1, 2)$ con $\rho_0 = 1000$





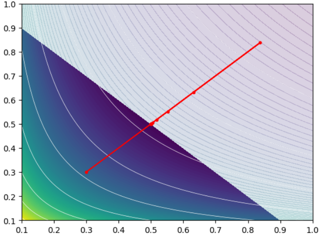
Nótese que se necesita un $\rho \approx 22,500$ para obtener una factibilidad aceptable $\|c(x)\| <$ 1e-6

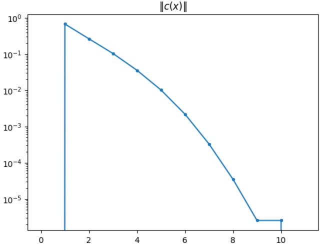
Ejemplo

Considérese $f(x, y) = -\ln x - \ln y$ con $x + y < 1$.
Utilizando valor inicial $(x, y) = (0.3, 0.3)$ con $\rho_0 = 1$

Ejemplo

Considérese $f(x, y) = -\ln x - \ln y$ con $x + y < 1$.
Utilizando valor inicial $(x, y) = (0.3, 0.3)$ con $\rho_0 = 1$





En este caso se necesitó un $\rho \approx 87$ para obtener $\|c(x)\| <$ 1e-6

Referencias

Michael Saunders, *Lecture Notes on Augmented Lagrangian Methods*, Stanford University, 2019. Available online at: https://stanford.edu/class/cme338/notes/notes11-BCL.pdf

María Daniela Sánchez, *Métodos de Lagrangiano Aumentado basados en funciones de penalidad no cuadráticas*, CONICET, 2017. Available online at: https://sedici.unlp.edu.ar/bitstream/handle/10915/65191/Documento_completo.pdf-PDFA.pdf?sequence=1