

## GLOBALIZACIÓN DE LA PROGRAMACIÓN CUADRÁTICA SUCESIVA CON REGIONES DE CONFIANZA

La principal desventaja de la programación secuencial cuadrática con búsqueda de línea es la exigencia de que  $B_k$  (aproximación del hessiano de la lagrangiana) sea definida positiva (exigencia poco natural, pues en el caso con restricciones el hessiano de la lagrangiana en el óptimo no tiene por que ser definido positivo). Una manera de trabajar con hessianos indefinidos en el caso sin restricciones son los métodos con región de confianza y la misma idea puede usarse para los problemas con restricciones.

En el caso de problemas con restricciones de igualdad:

$$\left\{ \begin{array}{l} \min_{x \in \mathbb{R}^n} f(x) \\ \text{verificando} \\ c(x) = 0, \quad \text{donde } c = (c_i)_{i=1}^{m_i} \end{array} \right.$$

la generalización del cálculo del paso con regiones de confianza es encontrar  $d_k$  solución de:

$$\left\{ \begin{array}{l} \min \nabla f(x_k)^t d + \frac{1}{2} d^t B_k d \\ c(x_k) + A(x_k) d = 0 \\ \|d\| \leq \Delta_k \end{array} \right. \quad (1)$$

Como no se exige que  $B_k$  sea definida positiva puede usarse  $B_k = \nabla_{xx}^2 \ell(x_k, \lambda_k)$ .

Hay una dificultad en esta elección: puede ocurrir que no exista solución de

$$\left\{ \begin{array}{l} c(x_k) + A(x_k) d = 0 \\ \|d\| \leq \Delta_k \end{array} \right. \quad (2)$$

Es necesario entonces encontrar alternativas. Vamos a presentar varias.

## 1 Método $S\ell_p$ QP (Fletcher)

La idea de este método es minimizar la función de penalización exacta  $\ell_p$ :

$$\Phi(x, \rho) = f(x) + \rho \|c(x)\|_p$$

para  $\rho$  suficientemente grande y alguna norma  $\ell_p$ . Por supuesto esta función de penalización es no diferenciable y será necesario construir un método especializado.

Como se trata de métodos con región de confianza, un modelo adecuado es el siguiente:

$$\begin{cases} \min \nabla f(x_k)^t d + \frac{1}{2} d^t B_k d + \rho \|c(x_k) + A(x_k) d\|_p \\ \|d\| \leq \Delta_k \end{cases} \quad (3)$$

Este problema es siempre consistente pues la única restricción es la cota sobre el paso.

## 2 Métodos de paso compuesto

Una alternativa para solventar las dificultades causadas por la inconsistencia de los subproblemas cuadráticos es separar el cálculo del paso en dos etapas. El objetivo de un método de paso compuesto es encontrar:

$$d_k = n_k + \tau_k$$

donde:

- el paso normal  $n_k$  se mueve hacia la parte admisible de las restricciones linealizadas (dentro de la región de confianza)
- el paso tangencial  $\tau_k$  reduce la función modelo sin sacrificar la admisibilidad obtenida con  $n_k$  (también dentro de la región de confianza)

El hecho de que el paso normal este relacionado únicamente con la admisibilidad hace que el objetivo modelo pueda empeorar y no recuperarse durante el paso tangencial. Como el paso tangencial tiene que mantener la admisibilidad (linealizada) adquirida en el paso normal se verifica que:

$$A(x_k)(n_k + \tau_k) = A(x_k)n_k$$

y por tanto  $A(x_k)\tau_k = 0$ .

Vemos ahora distintas elecciones de los pasos normal y tangencial.

### 2.1 Método de relajación de restricciones (Vardi)

El paso normal se obtiene relajando:

$$\begin{cases} c(x_k) + A(x_k)d = 0 \\ \|d\| \leq \Delta_k \end{cases} \quad (4)$$

a:

$$\begin{cases} \sigma_k c(x_k) + A(x_k)n = 0 \\ \|n\| \leq \Delta_k \end{cases} \quad (5)$$

donde  $\sigma_k$  es suficientemente pequeño para que exista un  $n_k$  admisible. Claramente  $d = 0$  es admisible si  $\sigma_k = 0$ , y el mayor  $\sigma_{max}$  posible puede encontrarse calculando:

$$\max_{\sigma \in (0,1]} \left( \min_{\|d\| \leq \delta_k} \|\sigma c(x_k) + A(x_k)d = 0\| \right)$$

En la práctica se elige algún valor entre 0 y  $\sigma_{max}$ . El principal defecto de este método es que puede no existir paso normal si las restricciones linealizadas son inconsistentes.

Una vez se ha determinado un paso normal, se calcula el paso tangencial como:

$$\begin{cases} \text{(aproximado) } \arg \min_{\tau \in \mathbb{R}^n} (\nabla f(x_k) + B_k n_k)^t \tau + \frac{1}{2} \tau^t B_k \tau \\ A(x_k) \tau = 0 \\ \|n_k + \tau\| \leq \Delta_k \end{cases} \quad (6)$$

## 2.2 Método de reducción de restricciones (Byrd-Omojokun)

Este método soluciona las dificultades relacionadas con la inconsistencia del método de Vardi. En lugar de relajar las restricciones, el paso normal se calcula como:

$$\begin{cases} \text{(aproximado) } \arg \min_{n \in \mathbb{R}^n} \|c(x_k) + A(x_k)n\| \\ \|n\| \leq \Delta_k \end{cases} \quad (7)$$

El paso tangencial se calcula exactamente igual que en el método de Vardi.

(Este método es la base del software KNITRO)

## 2.3 Método de acumulación de restricciones (Celis-Dennis-Tapia)

En este método las restricciones  $c(x_k) + A(x_k)d = 0$  se reemplazan por:

$$\|c(x_k) + A(x_k)d\| \leq \sigma_k$$

para algún  $\sigma_k \in [0, \|c(x_k)\|]$ . El valor de  $\sigma_k$  se elige de manera que el paso normal  $n_k$  verifique:

$$\|c(x_k) + A(x_k)n\| \leq \sigma_k \text{ y } \|n\| \leq \Delta_k$$

Una vez encontrado un paso normal adecuado, el paso tangencial se calcula como:

$$\begin{cases} \text{(aproximado) } \arg \min_{\tau \in \mathbb{R}^n} (\nabla f(x_k) + B_k n_k)^t \tau + \frac{1}{2} \tau^t B_k \tau \\ \|c(x_k) + A(x_k)n_k + A(x_k)\tau\| \leq \sigma_k \\ \|n_k + \tau\| \leq \Delta_k \end{cases} \quad (8)$$

Este último subproblema es muy complicado de resolver y el método apenas se usa en la práctica.

### 3 Métodos de filtro (Fletcher-Leyffer)

Al contrario que el resto de los métodos vistos hasta el momento, los métodos de filtro no usan una función de mérito para forzar la convergencia global.

- La principal objeción a las funciones de mérito es que dependen, en gran medida, de parámetros arbitrarios o desconocidos a priori.
- Una segunda objeción a las funciones de mérito es que son demasiado conservadoras aceptando potenciales iterantes

Si no se quieren usar funciones de mérito será necesario utilizar algún otro medio de forzar la convergencia. La nueva idea es usar **un filtro**.

Sea  $\theta(x) = \|c(x)\|$  alguna norma de la violación de las restricciones en  $x$ . Un **filtro** es un conjunto de pares  $\{(\theta_k, f_k)\}$  de violaciones y de valores de la función objetivo tales que ningún miembro domina al otro, es decir, no ocurre que

$$\theta_i < \theta_j \text{ y } f_i < f_j$$

para cualquier par de puntos del filtro con  $i \neq j$  (“ $<$ ” significa “muy ligeramente menor que”).

Cualquier potencial iterante  $x_k + d_k$  del método será rechazado si no es aceptable para el filtro acumulado en las  $k$  iteraciones anteriores. Esto puede introducirse en el marco de las regiones de confianza y una iteración típica sería como sigue:

- Si es posible, elegir

$$\begin{cases} d_k = \arg \min_{d \in \mathbb{R}^n} \nabla f(x_k)^t d + \frac{1}{2} d^t B_k d \\ A(x_k) d + c(x_k) = 0 \\ \|d\| \leq \Delta_k \end{cases} \quad (9)$$

- pero, en otro caso, buscar  $d_k$  tal que:

$$\theta(x_k + d_k) < \theta_i \text{ para todo } i \leq k$$

- Si  $x_k + d_k$  es aceptable para el filtro, hacer  $x_{k+1} = x_k + d_k$  y, posiblemente añadir  $(\theta(x_k + d_k), f(x_k + d_k))$  al filtro, podar el filtro e incrementar  $\Delta_k$ .
- En caso de que  $x_k + d_k$  no sea aceptable se reduce  $\Delta_k$  y se repite el proceso.