OPTIMIZACIÓN NUMÉRICA I
LABORATORIO
PUNTOS INTERIORES

Introducción

Sea $G \in \mathbb{R}^{n\times n}$ simétrica positiva (semi)definida, $A \in \mathbb{R}^{m\times n}, \vec{\boldsymbol{b}} \in \mathbb{R}^m, \vec{\boldsymbol{c}} \in \mathbb{R}^n$. Tenemos el problema

$$\begin{array}{ll} \text{minimizar} & \frac{1}{2}\vec{\boldsymbol{x}}^\top G\vec{\boldsymbol{x}} + \vec{\boldsymbol{c}}^\top \vec{\boldsymbol{x}} \\ \text{sujeto a} & A\vec{\boldsymbol{x}} \le \vec{\boldsymbol{b}} \end{array}$$

y con variables de holgura $\vec{\boldsymbol{y}}$:

$$\begin{array}{ll} \text{minimizar} & \frac{1}{2}\vec{\boldsymbol{x}}^\top G\vec{\boldsymbol{x}} + \vec{\boldsymbol{c}}^\top \vec{\boldsymbol{x}} \\ \text{sujeto a} & A\vec{\boldsymbol{x}} + \vec{\boldsymbol{y}} = \vec{\boldsymbol{b}} \end{array}$$

Las condiciones KKT (necesarias y suficientes) para el problema son:

$$F(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}, \vec{\boldsymbol{\mu}}) = \begin{pmatrix} G\vec{\boldsymbol{x}} + A^\top \vec{\boldsymbol{\mu}} + \vec{\boldsymbol{c}} \\ A\vec{\boldsymbol{x}} + \vec{\boldsymbol{y}} - \vec{\boldsymbol{b}} \\ D_y D_\mu \mathbb{1} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad (\vec{\boldsymbol{y}}, \vec{\boldsymbol{\mu}}) \ge \mathbf{0},$$

donde $D_y = \operatorname{diag}(\vec{\boldsymbol{y}})$, $D_\mu = \operatorname{diag}(\vec{\boldsymbol{\mu}})$ y $\mathbb{1} = (1, 1, \dots, 1)^\top \in \mathbb{R}^m$.

No conocemos una solución óptima aún.

Pero, si F' (el Jacobiano) es regular en (una) solución óptima $(\vec{\boldsymbol{x}}^\star, \vec{\boldsymbol{y}}^\star, \vec{\boldsymbol{\mu}}^\star)$, entonces (usando el Teorema de la función implícita) existe un intervalo $I = (\alpha, \beta)$ y un abierto $\mathcal{G} \subset \mathbb{R}^{n+m+m}$ que contiene a $(\vec{\boldsymbol{x}}^\star, \vec{\boldsymbol{y}}^\star, \vec{\boldsymbol{\mu}}^\star)$ tales que para $\tau \in I$ existe un único $(\vec{\boldsymbol{x}}_\tau, \vec{\boldsymbol{y}}_\tau, \vec{\boldsymbol{\mu}}_\tau)$ con

$$F(\vec{\boldsymbol{x}}_\tau, \vec{\boldsymbol{y}}_\tau, \vec{\boldsymbol{\mu}}_\tau) = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \tau\mathbb{1} \end{pmatrix}.$$

La trayectoria central es el conjunto

$$\mathcal{T}_c = \left\{ (\vec{\boldsymbol{x}}_\tau, \vec{\boldsymbol{y}}_\tau, \vec{\boldsymbol{\mu}}_\tau) \colon F(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}, \vec{\boldsymbol{\mu}}) = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \tau\mathbb{1} \end{pmatrix}, \ \tau \in (0, \beta) \right\}.$$

Dr. Andreas Wachtel, Primavera 2022.

Un método de Newton (sencillo).

Idea: Acercarse al óptimo seguiendo la trayectoria central.

Dado $0 < tol \ll 1$ y $\vec{\boldsymbol{w}}_0 = (\vec{\boldsymbol{x}}^0, \vec{\boldsymbol{y}}^0, \vec{\boldsymbol{\mu}}^0)$ con $(\vec{\boldsymbol{y}}^0, \vec{\boldsymbol{\mu}}^0) > \mathbf{0}$ definimos $k \coloneqq 0$ y $\sigma = 0.1$.

MIENTRAS $\|F(\vec{\boldsymbol{w}}_k)\|_\infty > tol$ & $k < 200$

- Calcule $M_k \coloneqq \frac{1}{m}(\vec{\boldsymbol{y}}^k)^\top \vec{\boldsymbol{\mu}}^k$ y resuelve

$$\begin{bmatrix} G & 0 & A^\top \\ A & I & 0 \\ 0 & D_\mu & D_y \end{bmatrix} \begin{bmatrix} \vec{\boldsymbol{\delta}}_x \\ \vec{\boldsymbol{\delta}}_y \\ \vec{\boldsymbol{\delta}}_\mu \end{bmatrix} = -F(\vec{\boldsymbol{w}}_k) + \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \sigma M_k \mathbb{1} \end{bmatrix}$$

- Determine los valores máximos $\alpha_y, \alpha_\mu \in (0, 1]$ tales que

$$\vec{\boldsymbol{y}}^k + \alpha_y \vec{\boldsymbol{\delta}}_y \geq 0 \quad , \quad \vec{\boldsymbol{\mu}}^k + \alpha_\mu \vec{\boldsymbol{\delta}}_\mu \geq 0\,.$$

  Consejo: Implementar una función que encuentre esos valores.

- Actualicé: Sea $\alpha \coloneqq \text{mín}\,\{\alpha_y, \alpha_\mu\}$ entonces

$$\vec{\boldsymbol{w}}_{k+1} \coloneqq \begin{bmatrix} \vec{\boldsymbol{x}}^{k+1} \\ \vec{\boldsymbol{y}}^{k+1} \\ \vec{\boldsymbol{\mu}}^{k+1} \end{bmatrix} \coloneqq \begin{bmatrix} \vec{\boldsymbol{x}}^{k} \\ \vec{\boldsymbol{y}}^{k} \\ \vec{\boldsymbol{\mu}}^{k} \end{bmatrix} + \frac{99}{100}\alpha \begin{bmatrix} \vec{\boldsymbol{\delta}}_x \\ \vec{\boldsymbol{\delta}}_y \\ \vec{\boldsymbol{\delta}}_\mu \end{bmatrix}$$

- Si $\alpha_y > 0.9$ y $\alpha_\mu > 0.9$, entonces la dirección de Newton coincide aproximadamente con la dirección de la trayectoria central. La constante $\sigma = 0.1$ limita la convergencia del método. Por lo tanto, en este caso es bueno reducir el valor de σ al máximo entre $\{10^{-4},\, \sigma/10\}$.

- $k \coloneqq k + 1$

FIN

Un problema de prueba.

Use $G = 10^{-8} I$ con la identidad $I \in \mathbb{R}^{n \times n}$ y el problema de Klee-Minty dado por:

$$\begin{array}{ll} \text{minimizar} & \frac{1}{2}\vec{\boldsymbol{x}}^\top G \vec{\boldsymbol{x}} - \sum_{i=1}^n x_i \\ \text{sujeto a} & x_1 \leq 1 \\ & 2\sum_{j=1}^{i-1} x_j + x_i \leq 2^i - 1 \quad , \; i = 2, \ldots, n \\ & x_1\,, \;\ldots\;, \; x_n \geq 0\,. \end{array}$$

Funciones en Python (MatLab). Escriba las siguientes funciones en Python:

```
function [alpha] = recorte( v, dv )
% In : v  ... vector > 0
%      dv ... vector
%
% Out: alpha ... máximo valor  0 < alpha <= 1  tal que  v + alpha*dv >= 0
```

Usando la función `recorte` implementa el algoritmo arriba:

```
function [xo, yo, muo, iter] = mPI_Newton(G, c, A, b)
% Do at most  200 iterations.
% In each iteration show the values: number of iteration, norm of F(wk)
%
% In : G ... nxn symmetric positive definite matrix
%      A ... mxn matrix
%      b ... column vector with as many rows as A
%      c ... column vector with as many columns as A
%
% Out: xo  ... optimal solution
%      yo  ... slack variables
%      muo ... Lagrange multipliers
%      iter ... number of iterations the method made.
%
% Parameters:
%   tol = 1e-9;  x0 = zeros(n,1);  y0 = ones(m,1);  mu0 = ones(m,1);
%
```