

A tandem queue with server slow-down and blocking

Mario Correddu

Università di Pisa

Metodi numerici per catene di Markov
prof. Beatrice Meini

August 10, 2023

Risolvere il problema *upper-level* :

$$\min_{x,y} F(x,y) \quad t.c. \quad G(x,y) \leq 0 \quad y \in S(x), \quad (1)$$

$S(x)$ definito a partire dal problema *lower-level*:

$$\min_z f(x,z) \quad t.c. \quad g(x,z) \leq 0$$

in cui si ha:

$$S(x) = \operatorname{argmin}_z \{f(x,z) | g(x,z) \leq 0\}$$

Risolvere il problema *upper-level* :

$$\min_{x,y} F(x,y) \quad t.c. \quad G(x,y) \leq 0 \quad y \in S(x), \quad (1)$$

$S(x)$ definito a partire dal problema *lower-level*:

$$\min_z f(x,z) \quad t.c. \quad g(x,z) \leq 0$$

in cui si ha:

$$S(x) = \operatorname{argmin}_z \{f(x,z) | g(x,z) \leq 0\}$$

Supponiamo $F, f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^q$ e
 $G : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^p$ siano di classe C^2

E

$$S(x) \neq \emptyset$$

Supponiamo $F, f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^q$ e $G : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^p$ siano di classe C^2

E

$$S(x) \neq \emptyset$$

Definiamo:

$$\varphi(x) := \min_z \{f(x, z) \mid g(x, z) \leq 0\} \quad (2)$$

Il nuovo problema diventa:

$$\min_{x,y} F(x, y) \quad t.c. \quad G(x, y) \leq 0 \quad g(x, y) \leq 0 \quad f(x, y) \leq \varphi(x)$$

- un algoritmo che permetta la risoluzione dei problemi a partire dalla riformulazione
- presentare delle proprietà che garantiscano la convergenza globale del metodo
- presentare delle proprietà che portino alla convergenza a un minimo locale stretto

Notazioni

Fissiamo i punti (\bar{x}, \bar{y}) e $(\bar{x}, \bar{z}) \in \mathbb{R}^{n+m}$:

Definiamo:

$$I^1 := I^G(\bar{x}, \bar{y}) := \{i \mid G_i(\bar{x}, \bar{y}) = 0\}$$

$$I^2 := I^g(\bar{x}, \bar{y}) \quad e \quad I^3 := I^g(\bar{x}, \bar{z})$$

$$\eta^1 := \eta^G(\bar{x}, \bar{y}, \bar{u}) := \{i \mid \bar{u}_i = 0, G_i(\bar{x}, \bar{y}) < 0\},$$

$$\theta^1 := \theta^G(\bar{x}, \bar{y}, \bar{u}) := \{i \mid \bar{u}_i = 0, G_i(\bar{x}, \bar{y}) = 0\},$$

$$\nu^1 := \nu^G(\bar{x}, \bar{y}, \bar{u}) := \{i \mid \bar{u}_i > 0, G_i(\bar{x}, \bar{y}) = 0\},$$

$$\eta^2 := \eta^g(\bar{x}, \bar{y}, \bar{v}), \quad \theta^2 := \theta^g(\bar{x}, \bar{y}, \bar{v}), \quad \nu^2 := \nu^g(\bar{x}, \bar{y}, \bar{v})$$

e

$$\eta^3 := \eta^g(\bar{x}, \bar{z}, \bar{w}), \quad \theta^3 := \theta^g(\bar{x}, \bar{z}, \bar{w}), \quad \nu^3 := \nu^g(\bar{x}, \bar{z}, \bar{w})$$

.

$$\min_{x,y} F(x,y) + \lambda(f(x,y) - \varphi(x)) \quad t.c. \quad G(x,y) \leq 0 \quad g(x,y) \leq 0 \quad (3)$$

Definition

il problema 1 è partially calm in uno dei suoi punti ammissibili (\bar{x}, \bar{y}) se esistono $\lambda \in (0, \infty)$ e un intorno U di $(\bar{x}, \bar{y}, 0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$ tali che:

$$F(x,y) - F(\bar{x}, \bar{y}) + \lambda|\zeta| \geq 0$$

per ogni $(x,y,\zeta) \in U$ con $G(x,y) \leq 0$, $g(x,y) \leq 0$,
 $f(x,y) - \varphi(x) + \zeta = 0$.

Theorem

Sia (\bar{x}, \bar{y}) un punto di minimo locale per il problema 1. Allora il problema è partially calm in (\bar{x}, \bar{y}) se e solo se esiste $\lambda \in (0, \infty)$ tale che (\bar{x}, \bar{y}) è un punto di ottimo per 3.

Definition

le condizioni upper-level di Mangasarian Fromovitz (UMFCQ) valgono, per un punto (\bar{x}, \bar{y}) , se esiste $d \in \mathbb{R}^{n+m}$, tale che:

$$\nabla G_i(\bar{x}, \bar{y})^T d < 0 \text{ per ogni } i \in I^1 \quad \text{e} \quad \nabla g_j(\bar{x}, \bar{y})^T d < 0 \text{ per ogni } j \in I^2$$

Definition

Dato \bar{x} vale la condizione lower-level di Mangasarian Fromovitz (LMFCQ) se per ogni $z \in S(\bar{x})$, esiste $d \in \mathbb{R}^m$ tale che :

$$\nabla g_j(\bar{x}, z)^T d < 0 \text{ per tutti } i, j \text{ vincoli attivi in } (\bar{x}, z)$$

Theorem

Sia (x, y) un punto di minimo locale per il problema 1, partially calm. Siano f e g_1, \dots, g_q convesse e supponiamo che valgano UMFCQ nel punto (\bar{x}, \bar{y}) e LMFCQ nel punto \bar{x} . Allora esistono $\lambda \in (0, \infty)$ e $u \in \mathbb{R}^p$, $(v, w) \in \mathbb{R}^q \times \mathbb{R}^q$ e $z \in \mathbb{R}^m$ tali che valgono :

$$\nabla_1 F(x, y) + \nabla_1 G(x, y)u + \nabla_1 g(x, y)v + \lambda \nabla_1 f(x, y) - \lambda \nabla_1 \ell(x, z, w) = 0 \quad (4)$$

$$\nabla_2 F(x, y) + \nabla_2 G(x, y)u + \nabla_2 g(x, y)v + \lambda \nabla_2 f(x, y) = 0 \quad (5)$$

$$\nabla_2 f(x, z) + \nabla_2 g(x, z)w = 0 \quad (6)$$

$$u \geq 0 \quad G(x, y) \leq 0 \quad u^T G(x, y) = 0 \quad (7)$$

$$v \geq 0 \quad g(x, y) \leq 0 \quad v^T g(x, y) = 0 \quad (8)$$

$$w \geq 0 \quad g(x, z) \leq 0 \quad w^T g(x, z) = 0 \quad (9)$$

- φ è convessa
- vale una versione generalizzata del teorema di Fritz John:

$$0 \in \nabla F(\bar{x}, \bar{y}) + \nabla G(\bar{x}, \bar{y})u + \nabla g(\bar{x}, \bar{y})v + \lambda \nabla f(\bar{x}, \bar{y}) + \lambda \begin{pmatrix} \partial(-\varphi)(\bar{x}) \\ \{0\} \end{pmatrix}$$

-

$$\partial\varphi(x) = \nabla_1 f(x, z) + \{w \nabla_1 g(x, z) \mid w \geq 0, w_i = 0 \text{ per i vincoli i inattivi}\}$$

Sia

$$\phi(a, b) = \sqrt{a^2 + b^2} - a - b$$

.

Riformulazione del sistema

Sia

$$\phi(a, b) = \sqrt{a^2 + b^2} - a - b$$

. Poniamo:

$$\phi^G(x, y, u) = \begin{pmatrix} \phi(-G_1(x, y), u_1) \\ \vdots \\ \phi(-G_p(x, y), u_p) \end{pmatrix} \quad e \quad \phi^g(x, y, v) = \begin{pmatrix} \phi(-g_1(x, y), v_1) \\ \vdots \\ \phi(-g_p(x, y), v_p) \end{pmatrix}$$

Sia

$$\phi(a, b) = \sqrt{a^2 + b^2} - a - b$$

. Poniamo:

$$\phi^G(x, y, u) = \begin{pmatrix} \phi(-G_1(x, y), u_1) \\ \vdots \\ \phi(-G_p(x, y), u_p) \end{pmatrix} \quad e \quad \phi^g(x, y, v) = \begin{pmatrix} \phi(-g_1(x, y), v_1) \\ \vdots \\ \phi(-g_p(x, y), v_p) \end{pmatrix}$$

e

$$\mathcal{L}^\lambda(x, y, u, v) = F(x, y) + u^T G(x, y) + v^T g(x, y) + \lambda f(x, y)$$

$$L^\lambda(x, y, z, u, v, w) = \mathcal{L}^\lambda(x, y, u, v) - \lambda \ell(x, z, w)$$

Riformulazione del sistema

Definiamo per $\zeta := (x, y, z, u, v, w)$:

$$\Phi^\lambda(\zeta) := \begin{bmatrix} \nabla L^\lambda(\zeta) \\ \phi^G(x, y, u) \\ \phi^g(x, y, v) \\ \phi^g(x, z, w) \end{bmatrix}$$

Allora il sistema del teorema è equivalente a chiedere $\Phi^\lambda(\zeta) = 0$

Definiamo per $\zeta := (x, y, z, u, v, w)$:

$$\Phi^\lambda(\zeta) := \begin{bmatrix} \nabla L^\lambda(\zeta) \\ \phi^G(x, y, u) \\ \phi^g(x, y, v) \\ \phi^g(x, z, w) \end{bmatrix}$$

Allora il sistema del teorema è equivalente a chiedere $\Phi^\lambda(\zeta) = 0$
Consideriamo ora:

$$\Psi^\lambda(\zeta) := \frac{1}{2} \|\Phi^\lambda(\zeta)\|^2$$

differenziabile con derivata continua.

Definition

Data $\psi : \mathbb{R}^n \rightarrow \mathbb{R}^m$, Lipschitz intorno a un punto \bar{x} lo Jacobiano generalizzato secondo Clarke è definito come :

$$\partial\psi(\bar{x}) := \text{conv}(\partial_B\psi(\bar{x}))$$

dove

$$\partial_B\psi(\bar{x}) := \left\{ \lim_{n \rightarrow \infty} \nabla\psi(x^n) \mid x^n \rightarrow \bar{x}, x^n \in D_\psi \right\}$$

Definition

Data $\psi : \mathbb{R}^n \rightarrow \mathbb{R}^m$, Lipschitz intorno a un punto \bar{x} , ψ è detta semismooth in \bar{x} se esiste per ogni $d \in \mathbb{R}^n$ il limite:

$$\lim_{\substack{V \in \partial\psi(\bar{x} + td') \\ d' \rightarrow d, \quad t \rightarrow 0}} Vd'$$

Algorithm Semismooth Newton algorithm

Passo 0: Si scelgono $\lambda > 0$, $\beta > 0$, $\epsilon \geq 0$, $t > 2$, $\rho \in (0, 1)$, $\sigma \in (0, \frac{1}{2})$, $\zeta_0 := (x^0, y^0, z^0, u^0, v^0, w^0)$.

Sia $k = 0$

Passo 1: while $\|\Psi^\lambda(\zeta^k)\| > \epsilon$ repeat:

Passo 2: Scegli $W^k \in \partial_B \Psi^\lambda(\zeta^k)$ e calcola la soluzione del sistema :

$$W^k d = -\Phi^\lambda(\zeta^k)$$

Se il sistema non ha soluzione o la condizione $\nabla \Psi^\lambda(\zeta^k)^T d^k \leq -\beta \|d^k\|^t$ è violata allora si pone semplicemente:

$$d^k := -\nabla \Psi^\lambda(\zeta^k)$$

Passo 3: trova il più piccolo intero s_k tale che

$$\Psi^\lambda(\zeta^k + \rho^{s_k} d^k) \leq \Psi^\lambda(\zeta^k) + \sigma \rho^{s_k} \nabla \Psi^\lambda(\zeta^k)^T d^k$$

Passo 4: Poni $\alpha_k = \rho^{s_k}$, $\zeta^{k+1} := \zeta^k + \alpha_k d^k$, $k := k + 1$

Definition

Sia $h : \mathbb{R}^k \rightarrow \mathbb{R}^k$. Diremo che h è BD-regolare nel punto x se ogni elemento di $\partial_B h(x)$ è non singolare. Similarmente h sarà CD-regolare in x se ogni elemento di $\partial h(x)$ è non singolare

Theorem

Valgono le seguenti:

- (a) Ogni punto di accumulazione della successione $\{\zeta_k\}$ generata dall'algoritmo è un punto stazionario di Ψ .*
- (b) Se uno dei punti limiti della successione $\{\zeta_k\}$, supponiamo ζ^* , è una soluzione BD-regolare del sistema $\Phi^\lambda = 0$, allora la convergenza è superlineare; se inoltre le F, G, f, g ammettono anche Hessiano Lipshitziano, allora la convergenza è quadratica.*

Definition

La qualificazione dei vincoli lower-level data da indipendenza lineare (LLICQ) è data da

$$\{\nabla_2 g_i(\bar{x}, \bar{z}) | g_i(\bar{x}, \bar{z}) = 0\}$$

è linearmente indipendente.

La qualificazione dei vincoli versione upper level di indipendenza lineare (ULICQ) è data dalla proprietà:

$$\{\nabla G_i(\bar{x}, \bar{y}) : i \in I_1\} \cup \{\nabla g_j(\bar{x}, \bar{y}) : j \in I_2\} \quad (10)$$

sono una famiglia linearmente indipendente

Definition

La qualificazione dei vincoli lower-level data da indipendenza lineare (LLICQ) è data da

$$\{\nabla_2 g_i(\bar{x}, \bar{z}) | g_i(\bar{x}, \bar{z}) = 0\}$$

è linearmente indipendente.

La qualificazione dei vincoli versione upper level di indipendenza lineare (ULICQ) è data dalla proprietà:

$$\{\nabla G_i(\bar{x}, \bar{y}) : i \in I_1\} \cup \{\nabla g_j(\bar{x}, \bar{y}) : j \in I_2\} \quad (10)$$

sono una famiglia linearmente indipendente

Il cono delle direzioni ammissibili dato da:

$$Q(\bar{x}, \bar{y}, \bar{z}) := \{d | \nabla G_i(\bar{x}, \bar{y})d^{1,2} = 0, \quad i \in \nu^1, \quad \nabla g_j(\bar{x}, \bar{y})d^{1,2} = 0, \\ j \in \nu^2, \quad \nabla g_j(\bar{x}, \bar{z})d^{1,3} = 0, \quad j \in \nu^3\}$$

Theorem

Sia $\bar{\zeta} := (\bar{x}, \bar{y}, \bar{z}, \bar{u}, \bar{v}, \bar{w})$, che soddisfa il sistema 4-9 per un certo $\lambda > 0$. Supponi che valga la versione upper level di qualificazione dei vincoli data da indipendenza lineare (ULICQ) e la qualificazione dei vincoli lower-level data da indipendenza lineare (LLICQ) nel punto (\bar{x}, \bar{z}) . Se inoltre:

$$(d^{1,2})^T \nabla^2 \mathcal{L}^\lambda(\bar{\zeta}) d^{1,2} > \lambda (d^{1,3})^T \nabla^2 \ell(\bar{\zeta}) d^{1,3} \quad (11)$$

per ogni $d \in Q(\bar{x}, \bar{y}, \bar{z}) \setminus \{0\}$ e $\theta^3(\bar{x}, \bar{z}, \bar{w}) = \emptyset$ allora Φ^λ è CD-regolare nel punto $\bar{\zeta}$.

Definition

Data una funzione $\psi : \mathbb{R}^n \rightarrow \mathbb{R}^m$, la sua derivata direzionale in \bar{x} nella direzione d è definita come:

$$\psi'(\bar{x}; d) = \lim_{t \downarrow 0} \frac{1}{t} (\psi(\bar{x} + td) - \psi(\bar{x}))$$

Definition

Data una funzione $\psi : \mathbb{R}^n \rightarrow \mathbb{R}^m$, la sua derivata direzionale seconda in \bar{x} nella direzione d ed e è definita come:

$$\psi''_{\lambda}(\bar{x}, \bar{y}; d; e) = \lim_{t \downarrow 0} \frac{1}{t^2} \left[\psi \left(\bar{x} + td + \frac{1}{2}t^2e \right) - \psi(\bar{x}) - t\psi'(\bar{x}; d) \right]$$

Theorem

sotto opportune ipotesi valgono le seguenti formule per le derivate direzionali di φ :

$$\varphi'(\bar{x}, d) = \min_{z \in S(\bar{x})} \nabla_1 \ell(\bar{x}, z, w_z)^T d$$

$$\varphi''(\bar{x}; d, e) = \min_{z \in S_1(\bar{x}; d)} \left\{ \nabla^2 \ell(\bar{x}, z, w_z) e + \xi_d(\bar{x}, z) \right\}$$

dove $S_1(\bar{x}; d) = \operatorname{argmin}_{z \in S(\bar{x})} \nabla_1 \ell(\bar{x}, z, w_z)^T d$, mentre $\xi_d(\bar{x}, z)$ è definito da

$$\begin{cases} \xi_d(\bar{x}, z) = \min_{e \in \mathcal{Z}_d(\bar{x}, z)} (d, e)^T \nabla^2 \ell(\bar{x}, z, w_z) (d, e) \\ \mathcal{Z}_d(\bar{x}, z) = \left\{ e \in \mathbb{R}^m \mid \nabla_1 g_i(\bar{x}, z)^T d + \nabla_2 g_i(\bar{x}, z)^T e, \quad i \in I^3 \right\} \end{cases}$$

Il cono delle direzioni ammissibili è dato da:

$$C^\lambda(\bar{x}, \bar{y}) = \{d \mid \nabla G_i(\bar{x}, \bar{y})^T d = 0 \text{ per } i \in \nu^1, \quad \nabla G_i(\bar{x}, \bar{y})^T d \leq 0 \text{ per } i \in \theta^1, \\ \nabla g_i(\bar{x}, \bar{y})^T d = 0 \text{ per } i \in \nu^2, \quad \nabla g_i(\bar{x}, \bar{y})^T d \leq 0 \text{ per } i \in \theta^2 \\ \nabla F(\bar{x}, \bar{y})^T d + \lambda f(\bar{x}, \bar{y})^T d - \lambda \nabla_1 \ell(\bar{x}, z, w)^T d^1 \leq 0 \text{ per } z \in S(\bar{x})\}$$

Il cono delle direzioni ammissibili è dato da:

$$C^\lambda(\bar{x}, \bar{y}) = \{d \mid \nabla G_i(\bar{x}, \bar{y})^T d = 0 \text{ per } i \in \nu^1, \quad \nabla G_i(\bar{x}, \bar{y})^T d \leq 0 \text{ per } i \in \theta^1, \\ \nabla g_i(\bar{x}, \bar{y})^T d = 0 \text{ per } i \in \nu^2, \quad \nabla g_i(\bar{x}, \bar{y})^T d \leq 0 \text{ per } i \in \theta^2 \\ \nabla F(\bar{x}, \bar{y})^T d + \lambda f(\bar{x}, \bar{y})^T d - \lambda \nabla_1 \ell(\bar{x}, z, w)^T d^1 \leq 0 \text{ per } z \in S(\bar{x})\}$$

La lagrangiana modificata:

$$\bar{\mathcal{L}}_\kappa^\lambda(x, y, u, w) = \kappa(F(x, y) + \lambda f(x, y)) + \sum_{i \in I^1(d)} u_i G_i(x, y) + \sum_{j \in I^2(d)} v_j g_j(x, y)$$

dove $I^1(d)$ e $I^2(d)$ sono gli indici appartenenti rispettivamente a I^1 e I^2 tali che $\nabla G_i(x, y)d = 0$ e $\nabla g_j(x, y)d = 0$.

Theorem

Sia $\zeta := (x, y, z, u, v, w)$ che soddisfa le condizioni 4-9 per qualche $\lambda > 0$. Supponi che il problema lower level sia convesso nel punto \bar{x} e che le condizioni di derivabilità valgano per tutti i $d \in C^\lambda(\bar{x}, \bar{y})$.

Allora (\bar{x}, \bar{y}) è un punto di ottimo locale stretto del problema 3 se per ogni $d \in C^\lambda(\bar{x}, \bar{y}) \setminus \{0\}$, esistono dei vettori u, v, z^t e κ_t , dove $z^t \in S_1(\bar{x}, d^1)$, per $t = 1, \dots, k$ per un certo k , tali che:

$$(d)^T \nabla \bar{\mathcal{L}}_{\kappa_0}^\lambda(\bar{x}, \bar{y}, u, w) d > \lambda \sum_{t=1}^k \kappa_t \xi_{d^1}(\bar{x}, z^t)$$

dove, $\xi_{d^1}(\bar{x}, z^t)$, sono quantità che dipendono dalla definizione di derivata seconda di φ , $\kappa_0 = \sum_{t=1}^k \kappa_t$

Condizione simil-Robinson

$$\nabla_1 \bar{\mathcal{L}}_{\kappa_0}^\lambda(\bar{x}, \bar{y}, u, v) - \lambda \sum_{t=1}^k \kappa_t \nabla_1 \ell(\bar{x}, z^t, w^t) = 0$$

$$\nabla_2 \bar{\mathcal{L}}_{\kappa_0}^\lambda(\bar{x}, \bar{y}, u, v) = 0$$

dove $\Lambda(\bar{x}, z^t) = \{w^t\}$, e

$$\kappa_0 + \sum_{i \in I^1(d)} u_i + \sum_{j \in I^2(d)} v_j = 1$$

$$\kappa_0 \geq 0, \quad u_i \geq 0 \text{ per } i \in I^1(d), \quad v_j \geq 0 \text{ per } j \in I^2(d)$$

Idea di dimostrazione

- ① ϕ_λ ammette derivata direzionale e vale

$$\phi'_\lambda(\bar{x}, \bar{y}; d) \geq 0 \quad \forall d \in \mathbb{R}^{n+m}$$

- ② ϕ_λ ammette derivata direzionale seconda e soddisfa la condizione:

$$\inf_{e \in \mathbb{R}^{n+m}} \phi''(\bar{x}, \bar{y}; d; e) > 0 \quad \text{per tutti i } d \neq 0 \quad \text{t.c.} \quad \phi'_\lambda(\bar{x}, \bar{y}; d) = 0$$

- ③ ϕ_λ soddisfa quella che viene chiamata condizione di epiregolarità del secondo ordine, ovvero:
per ogni $d \in \mathbb{R}^{n+m}$, $t \geq 0$, e ogni cammino $e : \mathbb{R}_+ \rightarrow \mathbb{R}^{n+m}$ tale che $te(t) \rightarrow 0$ per $t \downarrow 0$ vale:

$$\phi_\lambda \left((\bar{x}, \bar{y}) + td + \frac{1}{2} t^2 e(t) \right) \geq \phi_\lambda(\bar{x}, \bar{y}) + t \phi'_\lambda(\bar{x}, \bar{y}; d) + \frac{1}{2} t^2 \phi''_\lambda(\bar{x}, \bar{y}; d; e(t)) + o(t^2)$$

Dei problemi presenti nella libreria, i veri valori ottimi sono noti per il 70% di essi. Per ognuno dei valori del parametro di penalizzazione osserviamo che il metodo converge sempre almeno il 87 delle volte. Inoltre l'ordine di convergenza sperimentale è stato osservato essere oltre 1.5 per oltre il 70% dei problemi.

Numerical experiments

Grazie per l'attenzione

- [1] A. Fischer, A. B. Zemkoho, and S. Zhou, “Semismooth newton-type method for bilevel optimization: Global convergence and extensive numerical experiments,” *Optimization Methods and Software*, vol. 37, no. 5, pp. 1770–1804, 2022. doi: 10.1080/10556788.2021.1977810. eprint: <https://doi.org/10.1080/10556788.2021.1977810> (cit. on p. 1).