Projections aléatoires pour les problèmes quadratiques à contraintes quadratiques

Leo Liberti[1], Benedetto Manca[2], Pierre-Louis Poirion[3]

[1] LIX CNRS Ecole Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau, France
leo.liberti@polytechnique.edu

[2] Dipartimento di Matematica e Informatica, Università di Cagliari, Italy
bmanca@unica.it

[3] RIKEN Center for Advanced Intelligence Project, Tokyo, Japan
pierre-louis.poirion@riken.jp

Mots-clés : *Lemma de Johnson-Lindenstrauss, Programmation nonlinéaire.*

On considère la formulation de Programmation Mathématique (PM) suivante :

$$\left.\begin{array}{rrcl} \max\limits_{x\in\mathbb{R}^n} & x^\top Q^0 x + c^0\cdot x & & \\ \forall i\le m & x^\top Q^i x + c^i\cdot x & = & q_i \\ & Ax & \le & b, \end{array}\right\} \tag{1}$$

où les paramètres sont : les matrices symétriques $Q^0,\ldots,Q^m$, les vecteurs $c^0,\ldots,c^m\in\mathbb{R}^n$, la matrice A de dimensions $d\times n$, le vecteur $b\in\mathbb{R}^d$; et les variables de décision sont les composantes du vecteur $x\in\mathbb{R}^n$.

La forme (1) est bien générale : des inégalités quadratiques peuvent y être représentées par l'ajout des variables de "slack/surplus", et des équations linéaires s'obtiennent moyennant $Q^i=0$ pour les indices i appropriés. On indique par $\mathsf{val}(P)$ la valeur optimale du PM (1).

Le but de ce résumé est d'introduire une reformulation approximante de l'Éq. (1) de taille beaucoup plus petite, ce qui est utile pour résoudre des problèmes quadratiques de taille conséquente : on résout la version approchée, et on exploite sa solution pour construire une solution quasiment faisable de l'Éq. (1).

Plus précisément, pour m,n suffisamment grands il existe des entiers $k<r$ (avec $k\ll m$ et $r\ll n$, matrices aléatoires T (de taille $k\times m$) et R (de taille $r\times n$), un vecteur $\alpha\in\mathbb{R}^k_+$, et des matrices

$$\begin{array}{rrcl} & \bar{A} & = & AR^\top \\ \forall h\le k & \bar{Q}^h & = & \sum\limits_{i\le m} T_{hi} R Q^i R^\top \\ \forall h\le k & \bar{c}^h & = & \sum\limits_{i\le m} T_{hi} c^i R^\top, \end{array}$$

telles que la formulation de PM suivante :

$$\left.\begin{array}{rrcl} \max\limits_{u\in\mathbb{R}^r} & u^\top \bar{Q}^0 u + \bar{c}^0\cdot u & & \\ \forall h\le k & -\alpha_h \le u^\top \bar{Q}^h u + \bar{c}^h\cdot u - \sum\limits_{i\le m} T_{hi} q_i & \le & \alpha_h \\ & \bar{A}u & \le & b \end{array}\right\} \tag{2}$$

a des relations avec l'Éq. (1) (qu'on voulait résoudre) qui peuvent être exploités par rapport à l'admissibilité et à l'optimalité. Les principaux théorèmes qu'on veut prouver sont les suivants :

Théorème 1 (Erreur l'optimalité)**.** *Il y a des valeurs $\delta^L, \delta^U > 0$ telles que :*

$$\mathsf{val}(1) - \delta^L \le \mathsf{val}(2) \le \mathsf{val}(1) + \delta^U \tag{3}$$

avec une probabilité arbitrairement grande (PAG).

Par "PAG" on veut signifier l'existance d'une suite d'événements E_h tels que :

$$\mathbf{Prob}(\,E_h\,) \ge 1 - \mathsf{C}e^{-\mathsf{c}h}$$

pour des constantes C, c. Les constantes δ^L, δ^U dépendent de plusieurs paramètres liés à l'Éq. (2), y inclus les entiers k, r. Ceci fait si que l'Éq. (3) tient avec probabilité au moins $1 - \mathsf{C}e^{-\phi(k,r)}$, où ϕ est un polynôme de k, r.

Théorème 2 (Erreur d'admissibilité)**.** *Soit $\bar{u}$ une solution admissible de l'Éq. (2), et $\bar{x} = \bar{u}R$. S'il y a un $\theta > 0$ tel que $\mathbf{1}^\top \bar{x} \le \theta$, alors l'erreur ℓ_2 d'admissibilité des contraintes d'inégalité de l'Éq. (1) à $\bar{x}$ est bornée supérieurement par $O(\mathcal{M}(Q)\theta^2\sqrt{\ln n})$ avec PAG, où la fonction $\mathcal{M}(Q)$ dépend linéairement des normes ℓ_2 de certains "tranches" du tenseur Q.*

Ces théorèmes, qui ne sont pour le moment que des conjectures crédibles, peuvent être démontrés en utilisant certains résultats de [1] (optimalité) et en faisant une extension d'un résultat de [2] (admissibilité). On travaille couramment à des preuves plus complexes, mais qui amélioréraient les bornes δ^L, δ^U et les quantités PAG C, ϕ par rapport à celles qui s'obtiendraient en appliquant [1, 2].

Les résultats préliminaires ci-dessous sur des instances aléatoires nonconvexes soutiennent la crédibilité de nos conjectures.

p	m	k	n	d	dens	obj	feas	t_{org}	t_{prj}
210	50	15	100	30	0.1554	0.0	1e-4	1.96	**1.73**
210	50	15	100	30	0.1997	0.0	2e-4	1.88	**1.83**
210	50	15	100	30	0.2910	0.0	3e-4	7.01	**1.99**
420	100	30	200	60	0.1045	0.0	0.0	17.70	**13.60**
420	100	30	200	60	0.1522	0.0	0.0	36.45	**16.71**
420	100	30	200	60	0.2477	0.0	1e-4	43.17	**19.02**
630	150	45	300	90	0.0870	0.0	0.0	92.81	**77.43**
630	150	45	300	90	0.1352	0.0	0.0	174.78	**74.82**
630	150	45	300	90	0.2312	0.0	0.0	288.69	**91.15**
840	200	60	400	120	0.0776	0.0	0.0	311.89	**218.07**
840	200	60	400	120	0.1262	0.0	0.0	668.76	**186.01**
840	200	60	400	120	0.2228	0.0	0.0	930.14	**214.22**
1050	250	75	500	150	0.0720	0.0	0.0	814.66	**386.71**
1050	250	75	500	150	0.1204	0.0	0.0	1323.49	**460.38**
1050	250	75	500	150	0.2174	0.0	0.0	1895.02	**459.78**

Note : obj = $(f^* - \bar{f})/f^*$, feas = erreur maximum agrégé d'inadmissibilité

$t_{\text{prj}} = \sum \mathsf{CPU}$ pour : construire T, R ; projeter Q, c, A ; résoudre (2) ; reconstituer une solution de (1).

On a utilisé le solveur IPOPT pour résoudre (1) (qui donne la valeur f^*) et (2) (qui donne la solution $\bar{u}$) à partir d'un point de départ aléatoire. La reconstitution d'un optimum de (1) à partir de la solution de (2) se fait en résolvant (1) avec $\bar{u}R$ comme point de départ, pour obtenir la valeur optimale $\bar{f}$. Ainsi, quand notre technique de projection est utilisée avec un solveur local, elle se réduit au calcul d'un bon point de départ.

Le fait que nous obtenons toujours $f^* \approx \bar{f}$ pourrait s'expliquer avec le fait que nos instances aléatoires, bien que nonconvexes, soient néanmoins faciles à résoudre. Les écarts remarquables entre t_{org} et t_{prj} pourraient être dus à l'utilisation d'AMPL comme interface : on a obtenu des écarts plus limités avec PyOmo.

Références

[1] C. D'Ambrosio, L. Liberti, P.-L. Poirion, and K. Vu. Random projections for quadratic programs. *Mathematical Programming B*, 183 :619–647, 2020.

[2] L. Liberti, B. Manca, and P.-L. Poirion. Random projections for the distance geometry problem. In M. Noy et al., editor, *Proceedings of the workshop Discrete Mathematics Days*, Santander, 2022. Universidad de Cantabria.