On Linear Programs with Linear Complementarity Constraints

Jong-Shi Pang

Department of Industrial and Enterprise Systems Engineering

University of Illinois at Urbana-Champaign

presented at

The Second International Conference on Nonlinear Programming with Applications

Academy of Mathematics and Systems Science, Beijing China

Tuesday April 08, 2008, 09:15–10:25 AM

reporting joint work with John Mitchell and Jing Hu

Contents of Presentation

- Definition of an LPCC and goals of research
- Fundamental roles in mathematical programming
- An LPCC formulation of a general quadratic program **(new!)**
- The global resolution of the LPCC, via a logical Benders approach
- An application: a simply-bounded indefinite quadratic program
- Numerical results on bounded-variable quadratic programs
- Concluding remarks

Definition of an LPCC

Given: $c \in \mathbb{R}^n$, $d \in \mathbb{R}^m$, $e \in \mathbb{R}^m$, $f \in \mathbb{R}^k$, $A \in \mathbb{R}^{k\times n}$, $B \in \mathbb{R}^{k\times m}$, and $C \in \mathbb{R}^{k\times m}$.

Find $(x, y, w) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ in order to globally

$$\boxed{\begin{array}{ll} \underset{(x,y,w)}{\text{minimize}} & c^T x + d^T y + e^T w \\ \text{subject to} & Ax + By + Cw \geq f \\ \text{and} & 0 \leq y \perp w \geq 0, \end{array}}$$

where $a \perp b$ means that the two vectors are orthogonal; i.e., $a^T b = 0$.

Preliminary observations

An LPCC is equivalent to 2^m linear programs, each called a piece and derived from a subset $\alpha \subseteq \{1, \cdots, m\}$ with complement $\bar{\alpha}$:

$$\text{LP}(\alpha): \quad \boxed{\begin{array}{ll} \underset{(x,y,w)}{\text{minimize}} & c^T x + d^T y + e^T w \\ \text{subject to} & Ax + By + Cw \geq f \\ & w_\alpha \geq 0 = y_\alpha \\ \text{and} & w_{\bar{\alpha}} = 0 \leq y_{\bar{\alpha}} \end{array}}$$

Thus, there are 3 states of an LPCC in general:

- infeasibility–all pieces are infeasible
- unboundedness–one piece is feasible and unbounded below
- global solvability–objective is bounded below on all feasible pieces and at least one piece is feasible.

Goals

To develop a finite-time algorithm to resolve an LPCC in one of its 3 states, without complete enumeration of all the pieces and without any a priori assumptions and/or bounds.

To provide certificates for the respective states at termination:

- no infeasible piece, if LPCC is infeasible
- an unbounded piece, if LPCC is feasible but unbounded below
- a globally optimal solution, if it exists.

To leverage the state-of-the-art advances in linear and integer programming.

To apply the developed methodology broadly.

Fundamental importance

The LPCC plays the same important role in disjunctive nonlinear programs as a linear program does in convex programs. In addition to many applications of its own.

Novel paradigms in mathematical programming

- hierarchical optimization/equilibration
- inverse optimization/equilibration
- parameter identification/model validation in optimization/equilibration.

Key formulations for

- General quadratic programs
- B-stationary conditions of MPECs

– verification and computation without MPEC-constraint qualification

- global resolution of nonconvex quadratic programs.

Inverse Convex Quadratic Programming

Given: $Q \in \mathbb{R}^{n \times n}$ symmetric positive semidefinite, $A \in \mathbb{R}^{m \times n}$, $(\bar{x}, \bar{b}, \bar{c})$ in $\mathbb{R}^{n+m+n}$, a polyhedron $\Omega \subseteq \mathbb{R}^{n+m+n}$, and a polyhedral norm $\| \bullet \|$ on $\mathbb{R}^{n+m+n}$.

Find $(x, b, c) \in \mathbb{R}^{n+m+n}$ in order to

$$\boxed{\begin{array}{lll} \underset{(x,b,c)}{\text{minimize}} & \| (x, b, c) - (\bar{x}, \bar{b}, \bar{c}) \| & \\ \text{subject to} & (x, b, c) \in \Omega & \\ \text{and} & x \in \underset{x'}{\text{argmin}} & \frac{1}{2} (x')^T Q x' + c^T x' \\ & & \text{subject to } Ax' \leq b \text{ and } x' \geq 0 \end{array}}$$

where argmin = the set of minimizers of the lower-level optimization problem.

Rewriting the lower-level convex QP in terms of its equivalent KKT conditions yields an LPCC.

B-stationarity of MPCCs

Consider the mathematical program with complementarity conditions (Scheel and Scholtes 2000):

$$\boxed{\begin{array}{ll} \underset{z}{\text{minimize}} & \theta(z) \\ \text{subject to} & g(z) \leq 0 \\ \text{and} & 0 \leq G(z) \perp H(z) \geq 0 \end{array}}$$

A feasible solution z^* is B-stationary if (an LPCC in horizontal form):

$$\boxed{\begin{array}{lll} 0 \in & \underset{v}{\text{argmin}} & \theta(z^*) + \nabla\theta(z^*)^T v \\ & \text{subject to} & g(z^*) + Jg(z^*)v \leq 0 \\ & \text{and} & 0 \leq G(z^*) + JG(z^*)v \perp H(z^*) + JH(z^*)v \geq 0 \end{array}}$$

Nonconvex quadratic programming

Consider the nonconvex quadratic program:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \frac{1}{2} x^T Q x + c^T x \\ \text{subject to} & Ax \leq b, \end{array}$$

where Q is symmetric but <u>not</u> positive semidefinite.

On the set of stationary points,

$$\text{objective value} = c^T x - b^T \xi, \quad \text{for any KKT multiplier } \xi,$$

leading to the (equivalent??) LPCC:

$$\begin{array}{ll} \underset{(x,\xi)}{\text{minimize}} & c^T x - b^T \xi \\ \text{subject to} & 0 = c + Qx + A^T \xi \\ \text{and} & 0 \leq \xi \perp b - Ax \geq 0. \end{array}$$

(Giannessi-Tomasin 1973) If QP_{min} is finite, then $\text{QP}_{\text{min}} = \text{LPCC}_{\text{min}}$.
(Recall the classical result of Curtis Eaves for $\text{QP}_{\text{min}} > -\infty$.)

However, equivalence breaks down if $\text{QP}_{\text{min}} = -\infty$.
(Trivial counter-example: minimize $-x^2$.)

Therefore, is there an equivalent LPCC formulation in general?

The answer is yes! See next.

There is a strong need for the global resolution of an LPCC.

Known facts of a feasible QP

(Majthay71) A feasible vector x is a (strict) local minimum if and only if x is a KKT point and Q is (strictly) copositive on the critical cone of the QP at x.

(Eaves71) The QP attains a global minimum solution if and only if its objective function is bounded below on the feasible set, or equivalently, on the feasible rays; furthermore, this holds if and only if (a) Q is copositive on the recession cone of the feasible set, and (b) $(c + Qx)^T d \geq 0$ for all feasible vectors x and recession directions d satisfying $d^T Q d = 0$.

(Luo-Tseng92) The quadratic objective function attains finitely many values on the set of stationary points.

(Giannessi-Tomasin73) If the QP has a finite optimal solution, then the minimum objective value is equal to the minimum stationary values.

There is yet no finite test for the complete resolution of a QP.

The complete LPCC formulation

$$\underset{(x,d,\xi,\lambda,\mu,t,s)\in\mathbb{R}^{2n+3m+2}}{\text{minimize}} \quad -t$$

subject to

$0 = c + Qx + A^T\xi + t\,\mathbf{1}_n$	Lagrangian equation augmented by t
$0 = Qd + A^T\lambda - A^T\boxed{\mu} + s\,\mathbf{1}_n$	derived from a ray problem
$0 \leq \xi \perp b - Ax \geq 0$	standard complementarity
$0 \leq \mu \perp b - Ax \geq 0$	connecting ray condition with feasibility
$0 \leq \lambda \perp -Ad \geq 0$	ray complementarity I
$0 \leq \xi \perp -Ad \geq 0$	connecting KKT multiplier with ray
$0 \leq \mu \perp -Ad \geq 0$	ray complementarity II
$0 \leq s, \quad \mathbf{1}_n^T d \geq 1$	ensuring nonzero ray

assuming, without loss of generality, that $\{d : Ad \leq 0\} \subseteq \Re^n_+$; otherwise, write $x = x^+ - x^-$ with $x^\pm \geq 0$ and substitute throughout.

Sketch of derivation and equivalence

The truncated QP: for $\rho > 0$,

$$\underset{x}{\text{minimize}} \quad \tfrac{1}{2}\, x^T Q x + c^T x$$

$$\text{subject to} \quad Ax \le b \ \text{ and } \ \mathbf{1}^T x \le \rho.$$

The truncated homogeneous QP (copositivity test)

$$\underset{d}{\text{minimize}} \quad \tfrac{1}{2}\, d^T Q d$$

$$\text{subject to} \quad Ad \le 0 \ \text{ and } \ \mathbf{1}^T d = \rho.$$

Theorem. Suppose that the QP is feasible. This QP is unbounded below if and only if the LPCC has a feasible solution with a negative objective value.

Simplification under copositivity

Suppose that Q is copositive on the recession cone (checkable by solving an LPCC).

The auxiliary variable μ can be removed, resulting in

$$
\begin{array}{ll}
\underset{(x,d,\xi,\lambda,t,s)\in\mathbb{R}^{2n+2m+2}}{\text{minimize}} & -t \\
\text{subject to} & 0 \,=\, c + Qx + A^T\xi + t\,\mathbf{1}_n \\
& 0 \,=\, Qd + A^T\lambda + s\,\mathbf{1}_n \\
& 0 \,\leq\, \xi \,\perp\, b - Ax \,\geq\, 0 \\
& 0 \,\leq\, \lambda \,\perp\, -Ad \,\geq\, 0 \\
& 0 \,\leq\, \xi \,\perp\, -Ad \,\geq\, 0 \\
& 0 \,\leq\, s, \;\; \mathbf{1}_n^T d \,\geq\, 1.
\end{array}
$$

The global resolution of the LPCC

The LPCC:

$$\boxed{\begin{array}{ll} \underset{(x,y,w)}{\text{minimize}} & c^T x + d^T y + e^T w \\ \text{subject to} & Ax + By + Cw \geq f \\ \text{and} & 0 \leq y \perp w \geq 0 \end{array}}$$

Introducing a conceptually very large scalar $\theta > 0$,

$$\boxed{\begin{array}{llll} \underset{(x,y,w,z)}{\text{minimize}} & c^T x + d^T y + e^T w & & \\ \text{subject to} & Ax + By + Cw & \geq & f \\ & \theta\, z \geq w & \geq & 0 \\ & \theta(\, 1 - z\,) \geq y & \geq & 0 \\ \text{and} & z & \in & \{\, 0, 1\,\}^m \end{array}}$$

Deficiencies and a resolution

Applicable only to feasible LPCCs with bounded variables.

Checking feasibility is difficult, especially when $B \neq 0$.

Lastly, computing the bounds of the variables is time consuming, if theoretically doable

(think about bounding the dual variables of a lower-level LP in a bilevel linear program)

Can the scalar θ be treated implicitly, even if it does not exist?

(think about the 2-phase implementation of the big-M simplex method)

Toward a parameter-free IP formulation

For a binary $z \in \{0,1\}^m$ and a scalar $\theta > 0$, the LP$(\theta; z)$:

$$
\begin{array}{llcll}
\underset{(x,y,w)}{\text{minimize}} & c^T x + d^T y + e^T w & & & \\
\text{subject to} & Ax + By + Cw & \geq & f & (\lambda) \\
 & -w & \geq & -\theta\, z & (u) \\
 & -y & \geq & -\theta\,(\mathbf{1} - z) & (v) \\
\text{and} & w, y & \geq & 0, &
\end{array}
$$

and its dual DP$(\theta; z)$:

$$
\begin{array}{ll}
\underset{(\lambda,u^{\pm},v)}{\text{maximize}} & f^T\lambda - \theta\left[\, z^T u + (\,\mathbf{1} - z\,)^T v \,\right] \\
\text{subject to} & A^T\lambda = c \\
 & B^T\lambda - v \leq d \\
 & C^T\lambda - u \leq e \\
\text{and} & (\,\lambda, u, v\,) \geq 0,
\end{array}
$$

which is feasible if and only if $\exists\, \lambda \geq 0$ satisfying $A^T\lambda = c$.

The (un-parameterized) master LP

Given a binary z with $\alpha = \text{supp}(z)$ and complement $\bar{\alpha}$,

$$\begin{array}{|ll|} \hline \underset{\lambda}{\text{maximize}} & f^T\lambda \\ \text{subject to} & A^T\lambda = c \\ & (B^T\lambda)_{\bar{\alpha}} \le d \\ & (C^T\lambda)_{\alpha} \le e \\ \text{and} & \lambda \ge 0, \\ \hline \end{array}$$

obtained from DP($\theta; z$) by respecting the constraint

$$z^T u + (\mathbf{1} - z)v \le 0.$$

The master LP, which is dual to the primal LP(α) piece,

(a) has a finite optimal solution

(b) is feasible and unbounded, or

(c) is infeasible.

Logical Benders cuts: of the satisfiability kind

In case (a), let λ^p be an optimal solution, add the point cut:

$$\sum_{i \in \bar{\alpha} : (C^T \lambda^p - e)_i > 0} z_i + \sum_{i \in \alpha : (B^T \lambda^p - d)_i > 0} (1 - z_i) \geq 1$$

In case (b), let λ^r be an optimal solution, add the ray cut:

$$\sum_{i \in \bar{\alpha} : (C^T \lambda^r)_i > 0} z_i + \sum_{i \in \alpha : (B^T \lambda^r)_i > 0} (1 - z_i) \geq 1$$

In case (c), solve the homogeneous dual problem:

$$\boxed{\begin{array}{ll} \underset{\lambda}{\text{maximize}} & f^T \lambda \\ \text{subject to} & A^T \lambda = 0 \\ & (B^T \lambda)_{\bar{\alpha}} \leq 0 \\ & (C^T \lambda)_{\alpha} \leq 0 \\ \text{and} & \lambda \geq 0, \end{array}} : \left\{ \begin{array}{lll} \max = \infty & \Rightarrow & \text{valid ray cut} \\ \max = 0 & \Rightarrow & \textbf{unbounded LPCC} \end{array} \right.$$

The key steps in a finite algorithm

- Generate initial cuts by a problem-dependent pre-processing procedure.

- Solve a satisfiability feasibility system to determine a binary vector z with supp(z).

- Solve the primal/dual master LP(α) to obtain either a point or ray cut, or an unboundedness certificate; in the process, (improved) upper bounds to LPCC_{min} are obtained.

- Apply a problem-dependent procedure to sparsify the obtained cuts, by solving tight LP relaxations restricted by the sparsified cuts under testing, obtaining lower bounds to the LPCC_{min} in the process.

$$\sum_{i \in \mathcal{I}} z_i + \sum_{j \in \mathcal{J}} (1 - z_j) \geq 1$$

split into ($\mathcal{I}_1 \cup \mathcal{I}_2 \subseteq \mathcal{I}$ and $\mathcal{J}_1 \cup \mathcal{J}_2 \subseteq \mathcal{J}$):

$$\sum_{i \in \mathcal{I}_1} z_i + \sum_{j \in \mathcal{J}_1} (1 - z_j) \geq 1 \quad \text{and} \quad \sum_{i \in \mathcal{I}_2} z_i + \sum_{j \in \mathcal{J}_2} (1 - z_j) \geq 1;$$

- Add the sparsified cuts to update satisfiability system. Return.

An application: Simply-bounded indefinite QPs

Consider the nonconvex quadratic program:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \tfrac{1}{2}x^T Q x + c^T x \\ \text{subject to} & 0 \le x \le \mathbf{1}_n \end{array}$$

and the equivalent LPCC formulation:

$$\begin{array}{ll} \underset{(x,\xi)}{\text{minimize}} & c^T x - \mathbf{1}_n^T \xi \\ \text{subject to} & 0 \le x \perp c + Qx + \xi \ge 0 \\ \text{and} & 0 \le \xi \perp \mathbf{1}_n - x \ge 0. \end{array}$$

The conceptual IP:

$$\begin{array}{lll} \underset{(x,\xi,z,\lambda)}{\text{minimize}} & c^T x - \mathbf{1}_n^T \xi & \\ \text{subject to} & x \le \mathbf{1}_n - z, & 0 \le c + Qx + \xi \le \theta\, z \\ \text{and} & \xi \le \theta\, \lambda, & \mathbf{1}_n - x \le \mathbf{1} - \lambda \\ & x, \xi \ge 0, & z, \lambda \in \{0,1\}^n. \end{array}$$

Some ideas and key steps

• The logical cut: $\lambda + z \leq 1$, expressing x cannot equal 0 and 1 simultaneously.

• The second-order cuts: A cut of the satisfiability kind can be generated if the second-order necessary condition is violated at a stationary point.

• The master LP: Given binary z and λ with $\alpha = \text{supp}(z)$ and $\gamma = \text{supp}(\lambda)$ satisfying $\alpha \cap \gamma = \emptyset$ and with respective complements $\bar{\alpha}$ and $\bar{\gamma}$:

$$\begin{array}{lll}
\underset{x_i : i \in \bar{\alpha} \cap \bar{\gamma}}{\text{minimize}} & \displaystyle\sum_{k \in \bar{\alpha} \cap \bar{\gamma}} \bar{c}_k \, x_k & \\
\text{subject to} & \bar{c}_i + \displaystyle\sum_{k \in \bar{\alpha} \cap \bar{\gamma}} q_{ik} \, x_k \left\{ \begin{array}{c} \geq \\ = \\ \leq \end{array} \right\} 0, \quad i = 1, \ldots, n, & \\
\text{and} & 0 \leq x_k \leq 1, \quad k \in \bar{\alpha} \cap \bar{\gamma} &
\end{array}$$

where $\bar{c}_i \equiv c_i + \displaystyle\sum_{j \in \gamma} q_{ij}$. Solving this LP or its dual yields a point or ray cut.

• Local search to recover stationarity, occurring in sparsification.

• Convex second-order cone program relaxation, time consuming.

Numerical results compared with Vandenbussche-Nemhauser (2005)

50-vars; density 30% and 40%

iter	Time	VN Time	LPcnt	VN LPcnt	cnt_rx	cnt_dual	cnt_M	Gtime
1	4.23	(13.28)	262	(434)	259	0	3	0.27
40	13.36	(127.07)	978	(4825)	697	267	14	1.78
44	13.47	(87.91)	932	(2827)	690	229	13	0.05
53	43.28	(464.51)	1767	(11356)	1120	640	7	0.01
75	51.97	(455.61)	1927	(10561)	1205	712	10	0.34
35	42.33	(263.06)	1494	(6464)	978	511	5	0.01

iter	=	# of satisfiability IPs solved
Time	=	total time (in seconds), including verification of global optimality
LPcnt	=	cnt_rx + cnt_dual + cnt_M
cnt_rx	=	# of relaxed LPs solved in sparsification (lower bounding)
cnt_dual	=	# of homogeneous dual LPs solved in generation of ray cuts
cnt_M	=	# of master LPs solved in cut generation
Gtime	=	time global solution is found but global optimality is not verified
computer	=	Core Duo CPU 2.33 GHz 1.95 GB of RAM
computer	=	SUN ultra-80/2x450-MHz Ulta-SPARC-II proc. and 1-GB memory

40-variables; density 60% to 100%

iter	Time*	VN Time	LPcnt	VN LPcnt	cnt_rx	cnt_dual	cnt_M	Gtime
44	64.59	983.32	2063	20590	1480	573	10	0.08
10	10.34	14.42	473	568	442	25	6	0.06
1	15.02	10.09	403	350	400	0	3	0.06
94	79.20	229.41	2068	7622	1551	515	2	0.03
37	59.00	138	1574	4490	1182	386	6	0.08
20	28.14	26.86	809	1081	682	123	4	0.27
53	138.02	359.03	2484	9729	1740	740	4	0.02
42	82.00	515.81	1699	12690	1242	453	4	0.08
42	58.27	117.31	1149	3687	931	212	6	1.23
32	90.34	199.48	1454	6246	1077	372	5	0.06
38	111.16	241	1745	7202	1267	467	11	0.50
35	86.41	73.32	1408	2589	1093	312	3	0.01
38	183.97	263.61	2053	7190	1481	569	3	0.13
35	148.33	2169.81	1852	27928	1308	541	3	0.01
39	171.47	$\boxed{58.84\%}$	2096	36207	1460	628	8	7.01

* As of a few days ago, these times are reduced, some up to 25%.

Boxed problem is suboptimal obtained by VN after 4000 seconds.

Concluding remarks

- LPCCs are a fundamental class of nonconvex programs deserving full study.
- Presented a complete LPCC for a general indefinite QP.
- Sketched a parameter-free IP-based algorithm for the complete resolution of a general LPCC.
- Applied to the simply-bounded QP and described some novel cuts and ideas.
- Numerical results are superior to those of Nemhauser-Vandenbussche (2005).
- Currently comparing with Burer-Vandenbussche (2007) on large sized bounded-variable QPs.
- On-going work: extension, refinement, and application.

$\boxed{\textbf{Thank you!}}$