Complete Solutions to Mixed Integer Programming

Ning Ruan

School of Science, Information Technology and Engineering, University of Ballarat, Ballarat, Australia
Email: n.ruan@ballarat.edu.au

ABSTRACT

This paper considers a new canonical duality theory for solving mixed integer quadratic programming problem. It shows that this well-known NP-hard problem can be converted into concave maximization dual problems without duality gap. And the dual problems can be solved, under certain conditions, by polynomial algorithms.

1. Introduction

Mixed integer nonlinear programming refers to optimization problems which involve continuous and discrete variables [8]. In this paper, we consider the following constrained mixed integer quadratic programming:

$$(P_0) \quad \min P(x, y) = f(x) + c^T y \qquad (1)$$

$$\text{s.t. } g(x) + w^T y \le 0,$$

$$-1 \le y \le 1,$$
$$x \in R^n, y \in R^{n,}$$

where, $f(x) = 1/2 x^T A x$, $g(x) = 1/2 x^T B x - b x - d$, c, w, b are given vectors, d is a given scalar, and $A, B \succ 0$, $c < 0$. X_a is a feasible space defined by

$$X_a = \left\{ x \in R^n, y \in R^n | y_i \in \{-1,1\}, i = 1, \ldots, n \right\} \quad (2)$$

Problem of the form (1) has a broad spectrum of applications, including process industry (process design [2, 13, 18], production planning [14], supply chain optimization [1,3], logistics and so on), management science (scheduling problem), financial (portfolio optimization problems [22]), engineering (network design [23]), machine learning (semi-supervised support vector machines), and computational chemistry /biology (solvent design problems).

Various methods have been proposed for solving mixed integer programming, such as branch and bound [4,5,19,21,24], cutting plane, branch and cut [16], branch and reduce, outer approximation [6,7,15], hybrid methods, and penalty method [17]. But the difficulty for developing an efficient method for such mixed integer programming lies not only on the nonlinearity of the func-

tions involved, but also on existence of both discrete and continuous variables [20]. But if we introduce the canonical duality with some strategy, we can find global optima in polynomial time [10, 11, 12].

The rest of paper is arranged as follows. In section 2, we demonstrate how to rewrite the primal problem as a dual problem by using the canonical dual transformation. In section 3, optimality criterions for global solutions are discussed. Finally, in the last section, we present some conclusions.

2. Canonical Dual Transformation

Canonical duality theory [9] is a potentially powerful methodology which can be used to solve a large class of non-convex and discrete problems in nonlinear analysis, global optimization, and computational science.

Since $y \in \{-1,1\}^n$, one penalty term is added. Let a be a penalty factor, the original problem can be formulated

$$(P) \min P(x, y) = f(x) + c^T y + \frac{1}{2} a \left(y \circ y - e \right)^2$$

$$\text{s.t. } g(x) + w^T y \le 0$$

$$y \circ y - e = 0$$

$x \in R^n, y \in R^n$.

We choose the geometrically nonlinear operator

$$\xi = \Lambda(y) = y \circ y$$

then, the canonical function associated with this geometrical operator is

$$V(\xi) = \frac{1}{2} a (\xi - e)^2.$$

Let $\zeta \in R^n$ be the canonical dual variable corre-

sponding to ξ, we have

$$\zeta = \nabla V(\xi) = a(\xi - e),$$

And the Legendre conjugates of the function $V(\xi)$ defined by

$$V^*(\zeta) = \{\xi^T\zeta - V(\xi) : \zeta = \nabla V(\xi)\} = \frac{1}{2}a^{-1}\zeta^T\zeta.$$

Thus the total complementarily function can be defined by

$$\Xi(x,y,\varsigma,\sigma,\tau)$$
$$= f(x) + c^y + \xi^T\zeta - V(\xi) + \sigma(y\circledast y - e)$$
$$+ \tau(g(x) + w^T y) = 0.$$

By the criticality condition

$$\delta_y\Xi(x,y,\varsigma,\sigma,\tau) = 0,$$

We obtain

$$y = \frac{-(c+\tau w)}{2(\zeta+\sigma)}$$

Therefore, the canonical dual problem can be proposed as the following:

$$(P^d)\quad \max P^d(\varsigma,\sigma,\tau) \\ \text{s.t.}(\varsigma,\sigma,\tau) \in S_a \tag{3}$$

and

$$P^d(\varsigma,\sigma,\tau)$$
$$= -\frac{1}{2}\tau^2 b^T(A+\tau B)^{-1}b - \tau d \tag{4}$$
$$-\frac{1}{4}\frac{(c+\tau w)^2}{(\varsigma+\sigma)} - \frac{1}{2a}\varsigma^2 - e(\varsigma+\sigma),$$

where $a = 10$, e is a vector with all its entry 1. Its dual feasible space S_a is defined as

$$S_a = \{\varsigma \in R^n, \sigma \in R^n, \tau \in R | \tau \geq 0, \varsigma + \sigma \neq 0\}. \tag{5}$$

The notation $sta\{\cdot\}$ stands for finding all stationary points of $P^d(\varsigma,\sigma,\tau)$ over S_a. The following theorem shows that (P^d) is canonically (i.e., with zero duality gap) dual to the primal problem (P).

3. Global Optimality Condition

Theorem 1 The problem (P^d) is canonical dual to the primal problem (P) in the sense that if $(\bar{\varsigma},\bar{\sigma},\bar{\tau})$ is a KKT point of (P^d), then $(\bar{x},\bar{y})$ defined by

$$\bar{x} = \bar{\tau}(A+\bar{\tau}B)^{-1}b, \bar{y} = -\frac{c+\bar{\tau}w}{2(\bar{\varsigma}+\bar{\sigma})} \tag{6}$$

is a KKT point of (P), and

$$P(\bar{x},\bar{y}) = P^d(\bar{\varsigma},\bar{\sigma},\bar{\tau}) \tag{7}$$

Proof. By introducing a Lagrange multiplier

$$(\epsilon,\xi) \in R_-^n \times R_-^n\left(R_-^n := \{\epsilon \in R^n | \epsilon \leq 0\}\right),$$

the Lagrangian $L : S_a \times R_-^n \times R_-^n \to R$ associated with the problem (P^d) is

$$L(\varsigma,\sigma,\tau,\epsilon,\xi) = P^d(\varsigma,\sigma,\tau) - \epsilon^T\sigma - \xi^T\tau.$$

The criticality conditions

$$\nabla_\zeta L(\bar{\varsigma},\bar{\sigma},\bar{\tau},\epsilon,\xi) = 0,$$
$$\nabla_\sigma L(\bar{\varsigma},\bar{\sigma},\bar{\tau},\epsilon,\xi) = 0,$$
$$\nabla_\tau L(\bar{\varsigma},\bar{\sigma},\bar{\tau},\epsilon,\xi) = 0$$

lead to

$$\bar{\zeta} = a(\bar{y}\circledast\bar{y} - e), \tag{8}$$

$$\epsilon = \nabla_\sigma P^d(\bar{\varsigma},\bar{\sigma},\bar{\tau}) = \bar{y}\circ\bar{y} - \bar{v}, \tag{9}$$

$$\xi = \nabla_\tau P^d(\bar{\varsigma},\bar{\sigma},\bar{\tau}) = \bar{v}\circledast\bar{v} - \bar{v}, \tag{10}$$

and the KKT conditions

$$0 < \bar{\sigma} \perp \epsilon = 0, \tag{11}$$

$$0 < \bar{\tau} \perp \xi = 0, \tag{12}$$

where $y = 1/2\text{Diag}(\bar{\varsigma}+\bar{\sigma})^{-1}(c/2 - \bar{\tau}w)$, the notation $s\circledast t := (s_1 t_1, s_2 t_2, \ldots, s_n t_n)$ denotes the Hadamard product for any two vectors $s,t \in R^n$. This shows that if $(\bar{\varsigma},\bar{\sigma},\bar{\tau})$ is a KKT point of the problem (P^d), and then $(\bar{x},\bar{y})$ is a KKT point of the primal problem (P).

By using the equations (6) we have

$$\partial_{\bar{\varsigma}}P^d = \frac{(c+\bar{\tau}w)^2}{4(\bar{\varsigma}+\bar{\sigma})^2} - \frac{\bar{\varsigma}}{a} - e, \tag{13}$$

$$\partial_{\bar{\sigma}}P^d = \frac{(c+\bar{\tau})^2}{4(\bar{\varsigma}+\bar{\sigma})^2} = 0, \tag{14}$$

$$\partial_{\bar{\tau}}P^d = -\frac{(c+\bar{\tau}w)w}{2(\bar{\varsigma}+\bar{\sigma})} \leq 0, \tag{15}$$

and

$$\bar{\varsigma}(\bar{y}\circledast\bar{y} - e) = 0, \\ \bar{\tau}(g(\bar{x}) + w^T\bar{y}) - 0. \tag{16}$$

So, in terms of

$$\bar{x} = \bar{\tau}(A+\bar{\tau}B)^{-1}b, \bar{y} = -\frac{c+\bar{\tau}w}{2(\bar{\varsigma}+\bar{\sigma})}, \tag{17}$$

we have

$$P^d(\bar{\varsigma},\bar{\sigma},\bar{\tau}) = \frac{1}{2}\bar{x}^T A\bar{x} + \tau\left(\frac{1}{2}\bar{x}^T B\bar{x} - \bar{x}^T b - d\right)$$
$$-\frac{(c+\bar{\tau}w)^2}{4(\bar{\varsigma}+\bar{\sigma})^2} - \frac{1}{2a}\bar{\varsigma}^2 - e(\bar{\varsigma}+\bar{\sigma})$$

$$= f(\overline{x}) + \tau g(\overline{x}) - \frac{(c + \overline{\tau} w)^2}{4(\overline{\varsigma} + \overline{\sigma})^2} - \frac{1}{2a} \overline{\varsigma}^2$$

$$- e(\overline{\varsigma} + \overline{\sigma})$$

$$= f(\overline{x}) + c^T \overline{y} + (\overline{y} \circ \overline{y}) \overline{\varsigma} - \left(\frac{1}{2a} \overline{\varsigma}^2 + e\overline{\varsigma} \right)$$

$$+ \sigma(\overline{y} \circ \overline{y} - e) + \overline{\tau}(g(\overline{x}) + w^T \overline{y}).$$

From (13), we have

$$\overline{y} \circ \overline{y} = \frac{\overline{\varsigma}}{a} + e. \qquad (18)$$

Therefore,

$$(\overline{y} \circ \overline{y}) \overline{\varsigma} - \left(\frac{1}{2a} \overline{\varsigma}^2 + e\overline{\varsigma} \right) = \frac{1}{2} a (\overline{y} \circ \overline{y} - e)^2 . \quad (19)$$

Due to the fact that $y_i \in \{-1, 1\}, i = 1, \ldots, n,$ we have

$$P^d(\overline{\varsigma}, \overline{\sigma}, \overline{\tau}) = P(\overline{x}, \overline{y}). \qquad (20)$$

This proves the theorem.

This theorem shows that there is no duality gap between the primal problem and its canonical dual. In order to identify the global minimize, we need to introduce a useful feasible space

$$S_a^+ = \{ (\varsigma, \sigma, \tau) \in S_k | \varsigma + \sigma > 0 \} \qquad (21)$$

be a subset of S_a, and we have the following theorem.

Theorem 2 Suppose that the vector $(\overline{\varsigma}, \overline{\sigma}, \overline{\tau})$ is a critical point of the canonical dual function $P^d(\varsigma, \sigma, \tau)$. Let

$$\overline{x} = \overline{\tau}(A + \overline{\tau} B)^{-1} b, \quad \overline{y} = -\frac{(c + \overline{\tau} w)}{2(\overline{\varsigma} + \overline{\sigma})} \qquad (22)$$

If $(\overline{\varsigma}, \overline{\sigma}, \overline{\tau}) \in S_a^+$, then $(\overline{\varsigma}, \overline{\sigma}, \overline{\tau})$ is a global maximize of $P^d(\varsigma, \sigma, \tau)$ on S_a^+, the vector $(\overline{x}, \overline{y})$ is a global minimize of $P(x, y)$ on R^n, and

$$P(\overline{x}, \overline{y}) = \min_{(x,y) \in R^n \times R^n} \min_{(x,y) \in R^n \times R^n} P(x, y)$$

$$= \max_{(\varsigma, \sigma, \tau) \in S_a^+} \max_{(\varsigma, \sigma, \tau) \in S_a^+} P^d(\varsigma, \sigma, \tau) \quad (23)$$

$$= P^d(\overline{\varsigma}, \overline{\sigma}, \overline{\tau})$$

Proof. By Theorem 1 and the canonical duality theory, we know that vector $(\overline{\varsigma}, \overline{\sigma}, \overline{\tau}) \in S_a$ is a KKT point of the problem (P^d) if and only if $(\overline{x}, \overline{y})$ defined by

$$\overline{x} = \overline{\tau}(A + \overline{\tau} B)^{-1} b, \quad \overline{y} = -\frac{c + \overline{\tau} w}{2(\overline{\varsigma} + \overline{\sigma})}$$

is a critical point of the problem (P), and

$$P(\overline{x}, \overline{y}) = P^d(\overline{\varsigma}, \overline{\sigma}, \overline{\tau})$$

By the fact that the canonical dual function $P^d(\varsigma, \sigma, \tau)$ is concave on S_a^+, the critical point $(\overline{\varsigma}, \overline{\sigma}, \overline{\tau})$ is a

global maximize of $P^d(\varsigma, \sigma, \tau)$ over S_a^+, this proves the statement (23).

This theorem provides a sufficient condition for a global minimizer of the primal problem.

4. Conclusions

In this paper, the canonical duality theory has been applied to solve mixed integer programming problem. Theorems show that by the canonical dual transformation, primal problems can be converted into canonical dual problem. By the fact that the canonical dual function is concave on the dual feasible space, so the dual problem can be solved by well-developed deterministic optimization methods.

5. Acknowledgements

REFERENCES

[1] K. Aardal," Capacited Facility Location: Separation Algorithms and Computational Experience," *Mathematical Programming,* Vol. 81, No. 2, 1998, pp. 149-175. doi:10.1007/BF01581103

[2] A. Atamtürk, "Flow Pack Facets of the Single Node Fixed-charge Flow Polytope," *Operations Research Letters*, Vol. 29, No. 3, 2001, pp. 107-114. doi:10.1016/S0167-6377(01)00100-6

[3] I. Barany, T. J. Van Roy and L. A. Wolsey, "Strong Formulations for Multi-item Capacitated Lot Sizing," *Management Science*, Vol. 30, 1984, pp. 1255-1261. doi:10.1287/mnsc.30.10.1255

[4] P. Belotti, J. Lee, L. Liberti, F. Margot and A. Waechter, "Branching and Bounds Tightening Techniques for Non-convex MINLP," *Optimization Methods & Software*, Vol. 24, 2009, pp. 597-634. doi:10.1080/10556780903087124

[5] B. Borchers and J. E. Mitchell, "An Improved Branch and Bound Algorithm for Mixed Integer Nonlinear Programs," *Computer & Operations Research*, Vol. 21, No. 4, 1994, pp. 359-367. doi:10.1016/0305-0548(94)90024-8

[6] M. A. Duran and I. E. Grossmann, "An Outer-approximation Algorithm for a Class of Mixed-integer Nonlinear Programs," *Mathematical Programming*, Vol. 36, No. 3, 1986, pp. 307-339. doi:10.1007/BF02592064

[7] R. Fletcher and S. Leyffer, "Solving Mixed Integer Nonlinear Programs by Outer Approximation," *Mathematical Programming*, Vol. 66, No. 1, 1994, pp. 327-349. doi:10.1007/BF01581153

[8] C. A. Floudas, I. G. Akrotirianakis, S. Caratzoulas, C. A. Meyer and J. Kallrath, "Global Optimization in the 21st Century: Advances and Challenges," *Computers &*

Chemical Engineering, Vol. 29, 2005, pp. 1185-1202. doi:10.1016/j.compchemeng.2005.02.006

[9] D. Y. Gao, "Duality Principles in Nonconvex Systems: Theory, Methods and Applications," Kluwer Academic Publishers, Dordrecht/ Boston/ London, 2000. doi:10.1007/978-1-4757-3176-7

[10] D. Y. Gao and N. Ruan, "Solutions to Quadratic Minimization Problems with Box and Integer Constraints," *Journal of Global Optimization*, Vol. 47, No. 3, 2010, pp. 463-484. doi:10.1007/s10898-009-9469-0

[11] D. Y. Gao, N. Ruan and H. D. Sherali, "Solutions and Optimality Criteria for Nonconvex Constrained Global Optimization Problems with Connections between Canonical and Lagrangian Duality," *Journal of Global Optimization*, Vol. 45, No. 3, 2009, pp. 473-497.
doi:10.1007/s10898-009-9399-x

[12] D. Y. Gao, N. Ruan and H. D. Sherali, "Canonical Dual Solutions to Fixed Cost Quadratic Programs", In: A. Chinchuluun, P.M. Pardalos, R. Enkhbat and L. Tseveendorj, Eds., *Optimization and Optimal Control: Theory and Applications*, Springer, Vol. 39, 2010, pp. 139-156. doi:10.1007/978-0-387-89496-6_7

[13] F. Glover and H. D. Sherali, "Some Class of Valid Inequalities and Convex Hull Characterizations for Dynamic Fixed-charge Problems under Nested Condtraints," Vol. 40, No. 1, 2005, pp. 215-234.

[14] J. Kallrath, "Solving Planning and Design Problem in the Process Industry using Mixed-integer and Global Optimization," *Annals of Operations Research*, Vol. 140, No. 1, 2005, pp. 335-373. doi:10.1007/s10479-005-3976-2

[15] P. Kesavan, R. J. Allgor, E. P. Gatzke and P. I. Barton, "Outer Approximation Algorithms for Separable Nonconvex Mixed-integer Nonlinear Programs," *Mathematical Programming*, Vol. 1000, No. 3, 2004, pp. 517-535.

[16] P. Kesavan and P. I. Barton," Generalized Branch-and-cut Framework for Mixed-integer Nonlinear Optimization Problems," *Computers & Chemical Engineering*, Vol. 24, 2000, pp. 1361-1366.
doi:10.1016/S0098-1354(00)00421-X

[17] L. Grippo and S. Lucidi, "A Differentiable Exact Penalty Function for Bound Constrained Quadratic Programming Problems," *Optimization*, Vol. 22, No. 4, 1991, pp. 557-578. doi:10.1080/02331939108843700

[18] Z. Gu, G. L. Nemhauser and M. W. P. Savelsbergh, "Lifted Flow Cover Inequalities for Mixed 0-1 Integer Programs," *Math Program*, Vol. 85, No. 3, 1999, pp. 439-467. doi:10.1007/s101070050067

[19] O. K. Gupta and A. Ravindran, "Branch and Bound Experiments in Convex Nonlinear Integer Programming," *Management Science*, 1985, pp. 1533-1546.
doi:10.1287/mnsc.31.12.1533

[20] P. Hansen, B. Jaumard, M. Ruiz and J. Xiong, "Global Minimization of Indefinite Quadratic Functions Subject to Box Constraints," *Naval Research Logistics*, Vol. 40, 1993, pp. 373-392.
doi:10.1002/1520-6750(199304)40:3<373::AID-NAV32 20400307>3.0.CO;2-A

[21] S. Leyffer, "Integrating SQP and Branch-and-bound for Mixed Integer Nonlinear Programming," *Computational Optimization and Applications,* Vol. 18, No. 3, 2001, pp. 295-309. doi:10.1023/A:1011241421041

[22] X. Lin, C. A. Floudas and J. Kallarth, "Global Solution Approach for a Non-convex MINLP Problem in Product Portfolio Optimization.," *Journal of Global Optimization*, Vol. 32, No. 3, 2005, pp. 417-431.
doi:10.1007/s10898-004-5903-5

[23] M. W. Padberg, T. J. Van Roy and L. A. Wolsey, "Valid Linear Inequalities for Fixed Charge Problems," *Operations Research*, Vol. 33, 1985, pp. 842-861.
doi:10.1287/opre.33.4.842

[24] I. Quesada and I. E. Grossmann, "An IP/NLP based Branch and Bound Algorithm for Convex NINLP Optimization Problems," *Computers & Chemical Engineering*, Vol. 16, 1992, pp. 937-947.
doi:10.1016/0098-1354(92)80028-8