A Finite Branch-and-Bound Algorithm for Nonconvex Quadratic Programming via Semidefinite Programming

Samuel Burer[1] Dieter Vandenbussche[2]

[1]University of Iowa

[2]University of Illinois at Urbana-Champaign

Semidefinite Programming and Its Applications
Institute of Mathematical Sciences
National University of Singapore
January 12, 2006

Outline

Outline

Nonconvex Quadratic Programming

Definition

$$\max \ \frac{1}{2} x^T Q x + c^T x \qquad \text{(QP)}$$
$$Ax \le b$$
$$x \ge 0$$

- Poly-time solvable if $Q \preceq 0$ (Kozlov, Tarasov, Khachiyan, 1979)
- We assume $Q \not\preceq 0$ ... NP-hard (e.g., generalizes max stable set)
- Local methods (Gould and Toint, 2002)
- Relaxations
  - LP (Sherali and Tuncbilek, 1995)
  - SDP (...)
- Global methods (Pardalos, 1991; BARON)
  - Theoretically infinite branch-and-bound tree ... convergence in limit

Nonconvex QP Over the Box

Definition

$$\max \ \frac{1}{2} x^T Q x + c^T x \qquad \text{(BoxQP)}$$
$$0 \le x \le e$$

- Generalizes 0-1 QP (Rosenberg, 1972)
- Perhaps most studied special case of (QP)

Globally Solving Nonconvex QP

- Our goal is to compute a global solution of (QP)
  - We assume feasible region is nonempty and bounded with interior
- Would like a finite B&B tree
- Will also study (BoxQP)

Outline

First-Order KKT Conditions

Definition

A feasible point x is a first-order KKT point of (QP) if there exists $(y, z) \geq 0$ such that

$$\begin{aligned} A^T y - z &= Qx + c \\ (b - Ax) \circ y &= 0 \\ x \circ z &= 0 \end{aligned}$$

- An optimal x is necessarily a KKT point.

First-Order KKT Conditions (cont'd)

Theorem (Giannessi and Tomasin, 1973)

Suppose (x, y, z) *constitutes a KKT point. Then*

$$\frac{1}{2}x^T Qx + c^T x = \frac{1}{2}b^T y + \frac{1}{2}c^T x$$

Corollary

(QP) is equivalent to

$$\begin{aligned} \max \ & \frac{1}{2}b^T y + \frac{1}{2}c^T x \\ & Ax \le b \quad x \ge 0 \\ & A^T y - z = Qx + c \quad (y, z) \ge 0 \\ & (b - Ax) \circ y = 0 \quad x \circ z = 0 \end{aligned} \tag{KKT}$$

Finite Branching

Node of Enumeration Tree

$$
\begin{aligned}
\max \ & \frac{1}{2} b^T y + \frac{1}{2} c^T x \\
& Ax \le b \quad x \ge 0 \\
& A^T y - z = Qx + c \quad (y, z) \ge 0 \\
& (b - Ax) \circ y = 0 \quad x \circ z = 0
\end{aligned}
$$

Finite Branching

Node of Enumeration Tree

$$
\begin{aligned}
\max \quad & \frac{1}{2} b^T y + \frac{1}{2} c^T x \\
& Ax \le b \quad x \ge 0 \\
& A^T y - z = Qx + c \quad (y, z) \ge 0
\end{aligned}
$$

Finite Branching

Node of Enumeration Tree

$$\max \ \frac{1}{2}b^T y + \frac{1}{2}c^T x$$
$$Ax \le b \quad x \ge 0$$
$$A^T y - z = Qx + c \quad (y, z) \ge 0$$

$$x_j = 0 \quad (j \in F^x) \qquad z_j = 0 \quad (j \in F^z)$$
$$A_{i\cdot} x = b_i \quad (i \in F^{b-Ax}) \qquad y_i = 0 \quad (i \in F^y)$$

Finite Branching

Node of Enumeration Tree

$$
\begin{aligned}
\max \;\; & \frac{1}{2} b^T y + \frac{1}{2} c^T x \\
& Ax \le b \quad x \ge 0 \\
& A^T y - z = Qx + c \quad (y, z) \ge 0
\end{aligned}
$$

$$
\begin{array}{llll}
x_j = 0 & (j \in F^x) & z_j = 0 & (j \in F^z) \\
A_{i\cdot} x = b_i & (i \in F^{b-Ax}) & y_i = 0 & (i \in F^y)
\end{array}
$$

$$
\begin{array}{rl}
F^x, F^z \subseteq \{1, \ldots, n\} & F^x \cap F^z = \emptyset \\
F^{b-Ax}, F^y \subseteq \{1, \ldots, m\} & F^{b-Ax} \cap F^y = \emptyset
\end{array}
$$

Finite Branching

Node of Enumeration Tree

$$\max \quad \frac{1}{2}x^T Q x + c^T x$$
$$Ax \le b \quad x \ge 0$$
$$A^T y - z = Qx + c \quad (y,z) \ge 0$$

$$x_j = 0 \quad (j \in F^x) \qquad z_j = 0 \quad (j \in F^z)$$
$$A_{i\cdot} x = b_i \quad (i \in F^{b-Ax}) \qquad y_i = 0 \quad (i \in F^y)$$

$$F^x, F^z \subseteq \{1,\ldots,n\} \qquad F^x \cap F^z = \emptyset$$
$$F^{b-Ax}, F^y \subseteq \{1,\ldots,m\} \qquad F^{b-Ax} \cap F^y = \emptyset$$

Outline

Linear Programming Relaxations?

Each node in tree has natural LP relaxation. However...

Theorem (B, Vandenbussche)

The LP relaxation at the root node has unbounded objective value.

- Same will hold for many nodes in B&B tree
- Unboundedness due to the Lagrange multipliers (y, z)
- For the LP relaxation to be useful, we must somehow bound (y, z)

Bounding (y, z) in the Case of (BoxQP)

Theorem (Vandenbussche, Nemhauser (2005))

For (BoxQP), it holds that

$$y_i \le c_i + \sum_{j=1}^{n} \max(Q_{ij}, 0)$$

$$z_j \le -c_j - \sum_{i=1}^{n} \min(Q_{ij}, 0)$$

- They develop LP-based finite branch-and-cut for (BoxQP)
- Significantly outperforms other approaches for (BoxQP)
- Cutting very important as plain B&B was not competitive
  - Even with bounded (y, z), are plain LP relaxations enough?

Bounding (y, z) in General

- But for general (QP), the proposed LP-based B&B has the aforementioned *fatal flaw* of unboundedness
- Alternatives to LP relaxations?

- There exist direct SDP relaxations of (QP)
- Known to be quite strong in certain cases (Ye, 1999)
- Based on the "$X \succeq xx^T$" paradigm
  - $\frac{1}{2}x^TQx + c^Tx \;\leftrightarrow\; \frac{1}{2}Q \bullet X + c^Tx$
- Not appropriate for finite branching because do not involve (y, z)

- We will use SDP relaxations of (KKT), i.e., ones involving (y, z)
- For example, include linear constraint

$$\frac{1}{2}Q \bullet X + c^Tx = \frac{1}{2}b^Ty + \frac{1}{2}c^Tx$$

Bounding (y, z) in General (cont'd)

Theorem (B, Vandenbussche)

The multipliers (y, z) stay bounded with SDP relaxation (unlike with LP).

Thus, an SDP-based finite B&B approach for (QP) works!

- Need to check one subtle point [next slide]

Also:

- Even if LPs were bounded, would expect SDPs to give significantly tighter relaxations
  - Hopefully no (or less) need for cutting planes
- Actually more than one SDP relaxation to choose from

Fathoming Leaf Nodes with SDP

- For correctness, SDP relaxation should have no gap at leaf nodes
- Must establish $\frac{1}{2}Q \bullet X + c^T x = \frac{1}{2}x^T Qx + c^T x$ at leaf
  - Not obvious because likely that $X \neq xx^T$

Proposition (B, Vandenbussche)

At a leaf node, the SDP relaxation is tight.

Corollary

B&B based on the SDP relaxation is correct (and finite).

Outline

Direct SDP Relaxation

Inspired by Lovász-Schrijver...

Definitions

$$Y := \begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} \in \Re^{(1+n)\times(1+n)}$$

$$K := \left\{ (x_0, x) \in \Re^{1+n} \; : \; Ax \le x_0\, b \quad (x_0, x) \ge 0 \right\}$$

$$\tilde{Q} := \frac{1}{2} \begin{pmatrix} 0 & c^T \\ c & Q \end{pmatrix}$$

Implications

$$Y e_0 = (1; x)$$

$$Y e_i \in K \quad \forall \quad i = 1, \ldots, n$$

$$Y \succeq 0$$

$$\tilde{Q} \bullet Y = \frac{1}{2}\, x^T Q x + c^T x$$

Direct SDP Relaxation (cont'd)

By relaxing the rank-1 condition:

(SDP$_0$)

$$
\begin{array}{rl}
\max & \tilde{Q} \bullet Y \\
\text{s.t.} & Ax \le b \quad x \ge 0 \\
& Ye_0 = (1; x) \\
& Ye_i \in K \quad \forall \quad i = 1, \ldots, n \\
& Y \succeq 0
\end{array}
$$

Note. Not suitable for finite branching because no way to enforce complementarities. However, by adding (y, z) ...

KKT SDP Relaxation

(SDP$_1$)

$$
\begin{aligned}
\max \quad & \tilde{Q} \bullet Y \\
\text{s. t.} \quad & Ax \le b \quad x \ge 0 \\
& Y e_0 = (1; x) \\
& Y e_i \in K \quad \forall \quad i = 1, \ldots, n \\
& Y \succeq 0 \\
& A^T y - z = Qx + c \quad (y, z) \ge 0 \\
& \tilde{Q} \bullet Y = \frac{1}{2} b^T y + \frac{1}{2} c^T x
\end{aligned}
$$

Strengthened KKT SDP Relaxation

Definitions

$$\hat{Y} := \begin{pmatrix} 1 \\ x \\ y \\ z \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \\ z \end{pmatrix}^T \in \Re^{(1+n+m+n)\times(1+n+m+n)}$$

$$\hat{K} := \text{homogenization of } \left\{ (x,y,z) \geq 0 : \begin{array}{c} Ax \leq b \\ A^T y - z = Qx + c \end{array} \right\}$$

$$\hat{Q} := \frac{1}{2} \begin{pmatrix} 0 & c^T & 0 & 0 \\ c & Q & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Strengthened KKT SDP Relaxation (cont'd)

Implications

$$
\begin{aligned}
\hat{Y}e_0 &= (1; x; y; z) \\
\hat{Y}e_i &\in \hat{K} \quad \forall \quad i = 1, \ldots, n+m+n \\
\hat{Y} &\succeq 0 \\
\hat{Q} \bullet \hat{Y} &= \frac{1}{2} x^T Q x + c^T x \\
(b - Ax) \circ y &= 0 \\
x \circ z &= 0
\end{aligned}
$$

Strengthened KKT SDP Relaxation (cont'd)

Implications

$$
\begin{aligned}
\hat{Y} e_0 &= (1; x; y; z) \\
\hat{Y} e_i &\in \hat{K} \quad \forall \quad i = 1, \ldots, n + m + n \\
\hat{Y} &\succeq 0 \\
\hat{Q} \bullet \hat{Y} &= \frac{1}{2} x^T Q x + c^T x \\
\mathrm{diag}(A \hat{Y}_{xy}) &= b \circ y \\
\mathrm{diag}(\hat{Y}_{xz}) &= 0
\end{aligned}
$$

Strengthened KKT SDP Relaxation (cont'd)

By relaxing the rank-1 condition:

(SDP$_2$)

$$
\begin{aligned}
\max \quad & \hat{Q} \bullet \hat{Y} \\
\text{s.t.} \quad & Ax \le b \quad A^T y - z = Qx + c \quad (x, y, z) \ge 0 \\
& \hat{Y} e_0 = (1; x; y; z) \\
& \hat{Y} e_i \in \hat{K} \quad \forall \quad i = 1, \ldots, n \\
& \hat{Y} \succeq 0 \\
& \tilde{Q} \bullet Y = \frac{1}{2} b^T y + \frac{1}{2} c^T x \\
& \mathsf{diag}(A \hat{Y}_{xy}) = b \circ y \\
& \mathsf{diag}(\hat{Y}_{xz}) = 0
\end{aligned}
$$

Outline

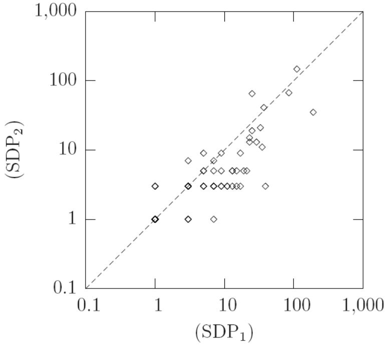
Solving the SDPs

- In our experiments, we employ
  - (SDP_1) and (SDP_2) for general (QP)
  - A single, streamlined SDP for (BoxQP)
- These SDPs are too large for standard codes
- Thus, use an augmented Lagrangian method (B, Vandenbussche)
- Additional advantages:
  - Dual feasible method $\Longrightarrow$ valid bounds at all times
  - Amenable to hot-starts

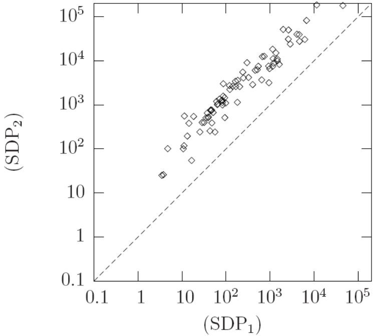
(QP): Testing Environment

- 78 randomly generated instances of (QP)
- Between 10 to 60 variables
- Between 10 to 60 constraints (not including $x \geq 0$)
- Compare performance of the two SDPs, (SDP$_1$) and (SDP$_2$)
- 2.4 GHz, 1 GB, Linux, C

(QP): No Optimality Tolerance

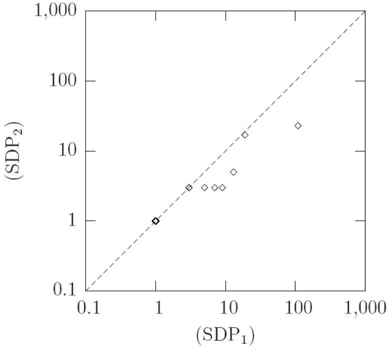


Number of Nodes

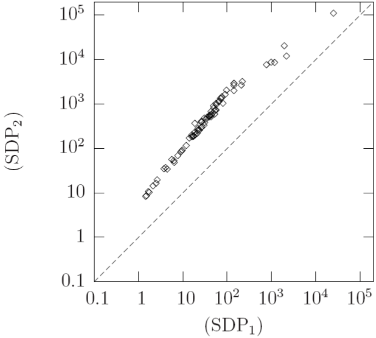


CPU Times

(QP): 1% Optimality Tolerance



Number of Nodes

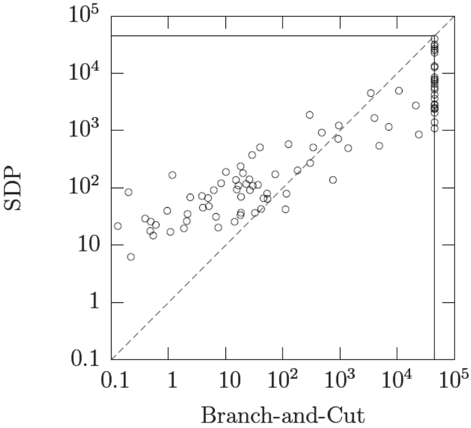


CPU Times

(BoxQP): Testing Environment

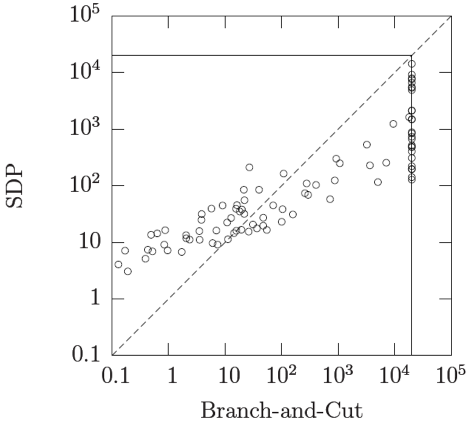
- 90 randomly generated instances of (BoxQP)
  - Between 20 and 100 variables
  - Smallest 54 instances (size 20 to 60) taken from Vandenbussche-Nemhauser
  - Largest 36 instances (size 70 to 100) newly generated
- Compare performance of SDP with LP-based branch-and-cut

(BoxQP): No Optimality Tolerance



CPU Times

(BoxQP): 1% Optimality Tolerance



CPU Times

Conclusions

- For general (QP)
  - Finite branching accomplished through complementarity
  - But one must bound (y, z) ... SDP to the rescue!
  - SDP relaxations very tight ... cutting planes less critical
- For (BoxQP), the SDP approach
  - Outperforms existing methods, including LP-based finite branch-and-cut method
  - Enables the solution of largest problems to date

Conclusions

- For general (QP)
  - Finite branching accomplished through complementarity
  - But one must bound (y, z) . . . SDP to the rescue!
  - SDP relaxations very tight . . . cutting planes less critical
- For (BoxQP), the SDP approach
  - Outperforms existing methods, including LP-based finite branch-and-cut method
  - Enables the solution of largest problems to date

THANK YOU!