Convexification of Mixed-Integer Quadratically Constrained Quadratic Programs

Laura Galli[1] Adam N. Letchford[2]

Lancaster, April 2011

[1]DEIS, University of Bologna, Italy
[2]Department of Management Science, Lancaster University, United Kingdom

Outline

- Literature Review
  - Motivation
  - Relaxations of QCQPs
  - Quadratic Convex Reformulation (QCR)
- Convexification of 0-1 QCQP
  - Reformulation with equations only
  - Reformulation with inequalities
- Convexification of MIQCQP
  - General integer variables
  - Continuous variables
- Conclusions

Motivation

It is well known that Semidefinite Programming (SDP) can be used to form useful relaxations for a variety of $\mathcal{NP}$-hard optimisation problems, such as:

- zero-one linear programs (0-1 LP)
- zero-one quadratic programs (0-1 QP)
- non-convex quadratically constrained quadratic programs (QCQP)
- general polynomial programs

Motivation (cont.)

Recently, Billionnet *et al.* (2009, 2010) proposed to use SDP to reformulate 0-1 QPs, not just to relax them.

The idea is to perturb the objective function in such a way that it is convex and the continuous relaxation of the 0-1 QP is tighter.

We examine to what extent this approach can be applied to the (much) more general case of mixed-integer quadratically constrained quadratic programs (MIQCQP).

Semidefinite relaxation of QCQP

We start with Shor's (1987) relaxation of non-convex QCQP (which is an $\mathcal{NP}$-hard global optimisation problem).

An instance takes the form:

$$\begin{array}{ll} \inf & x^T Q^0 x + c^0 \cdot x \\ \text{s.t.} & x^T Q^j x + c^j \cdot x \leq b_j \quad (j = 1, \ldots, m) \\ & x \in \mathbb{R}^n, \end{array}$$

where the Q^j are symmetric, square and rational matrices of order n, the c^j are rational n-vectors, and the b_j are rational scalars.

Semidefinite relaxation of QCQP (cont.)

The basic SDP relaxation of the QCQP is derived as follows:

- define the $n \times n$ symmetric matrix $X = xx^T$
- replace the terms $x^T Q^j x$ with $Q^j \bullet X$
- define the augmented symmetric matrix

$$Y = \begin{pmatrix} 1 \\ x \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}^T = \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}$$

- note that Y is positive semidefinite (psd).

Semidefinite relaxation of QCQP (cont.)

The resulting SDP is:

$$
\begin{array}{rl}
\inf & Q^0 \bullet X + c^0 \cdot x \\
\text{s.t.} & Q^j \bullet X + c^j \cdot x \leq b_j \quad (j = 1, \ldots, m) \\
& Y \succeq 0.
\end{array}
$$

(The last constraint imposes psd-ness on Y.)

Lagrangian relaxation of QCQP

As noted by Fujie & Kojima (1997), Lémaréchal & Oustry (2001) and others, there is a connection between semidefinite relaxation and Lagrangian relaxation.

Suppose we relax all of the constraint in Lagrangian fashion, and let λ denote the vector of Lagrangian multipliers.

The Lagrangian objective becomes:

$$x^T \left(Q^0 + \sum_{j=1}^{m} \lambda_j Q^j \right) x + \left(c^0 + \sum_{j=1}^{m} \lambda_j c^j \right) \cdot x - \sum_{j=1}^{m} \lambda_j b_j .$$

Lagrangian relaxation of QCQP (cont.)

Let the Lagrangian objective be denoted by $f(x, \lambda)$.

The relaxed problem is:

$$\inf \left\{ f(x, \lambda) : \; x \in \mathbb{R}^n \right\},$$

which is an unconstrained quadratic minimisation problem.

The Lagrangian dual is:

$$\sup_{\lambda \in \mathbb{R}^m_+} \inf_{x \in \mathbb{R}^n} f(x, \lambda).$$

Lagrangian relaxation of QCQP (cont.)

Suppose that strong duality holds for the SDP.

Then one can show that:

- the value of the Lagrangian dual equals the SDP bound,
- and the optimal Lagrangian multiplier vector λ is nothing but the optimal dual vector for the SDP.

QCQP summary

We have:

$$\text{opt} \geq \text{SDP bound} \geq \text{dual SDP bound} = \text{Lagrangian bound}$$

and the second inequality is an equality under suitable constraint qualifications.

0-1 QP

Now consider a 0-1 QP of the form:

$$\begin{aligned} \min \quad & x^T Qx + c^T x \\ \text{s.t.} \quad & Ax = b \\ & Dx \leq f \\ & x \in \{0,1\}^n. \end{aligned}$$

As observed by many authors, the condition that x be binary is equivalent to requiring

$$x_i - x_i^2 = 0 \qquad (i = 1, \ldots, n).$$

SDP relaxation of 0-1 QP

This leads immediately to the following semidefinite relaxation:

$$\begin{aligned} \min \quad & Q \bullet X + c^T x \\ \text{s.t.} \quad & Ax = b \\ & Dx \leq f \\ & x = \operatorname{diag}(X) \\ & Y \succeq 0. \end{aligned}$$

Convex reformulation of 0-1 QP

Reformulating a 0-1 QP instance means perturbing the objective function in such a way that:

- the objective function is convex
- the cost of each feasible solution is unchanged
- the constraints remain unchanged, so that the feasible set is unchanged.

Minimum eigenvalue reformulation of 0-1 QP

A simple reformulation technique was given by Hammer & Rubin (1970).

They add $\sum_{1 \leq i \leq n} M(x_i^2 - x_i)$ to the objective function, where $M > 0$ is large enough to ensure that the matrix $Q + MI$ of quadratic coefficients is psd.

In this way, they ensure that the reformulated instance is convex.

(A suitable value for M is the minimum eigenvalue of Q, multiplied by -1.)

Quadratic Convex Reformulation (QCR)

Billionnet *et al.* (2009) develop this idea further, in two ways:

- they perturb the objective function by addinggeneral terms of the form $\lambda_i(x_i - x_i^2)$ and $\mu_{jk}(a^j \cdot x - b_j)x_k$.
- instead of merely ensuring that the reformulated instance is convex, they also ensure that its continuous relaxation is as tight as possible.

Quadratic Convex Reformulation (QCR) (cont.)

Billionnet *et al.* (2009) show that the optimal multipliers λ and μ can be found by solving an SDP.

They call this approach convex quadratic reformulation or QCR.

Once QCR has been applied, one can feed the reformulated instance into any convex MIQP solver.

(An extension of QCR to general MIQP was given recently by Billionnet *et al.*, 2010.)

0-1 QCQP

Our first goal is to extend the QCR method to 0-1 QCQP.

We suppose the instance is written as:

$$\begin{array}{rll} \min & x^T Q^0 x + c^0 \cdot x & \\ \text{s.t.} & x^T Q^j x + c^j \cdot x = b_j & (j \in N_1) \\ & x^T Q^j x + c^j \cdot x \leq b_j & (j \in N_2) \\ & x \in \{0,1\}^n. & \end{array}$$

(We assume for simplicity of notation that any linear constraints, along with any additional valid inequalities, are subsumed in the above system.)

SDP relaxation of 0-1 QCQP

The semidefinite relaxation is immediate:

$$\begin{array}{rl} \min & Q^0 \bullet X + c^0 \cdot x \\ \text{s.t.} & Q^j \bullet X + c^j \cdot x = b_j \quad (j \in N_1) \\ & Q^j \bullet X + c^j \cdot x \le b_j \quad (j \in N_2) \\ & x = \operatorname{diag}(X) \\ & Y \succeq 0. \end{array}$$

The Lagrangian relaxation is easy to write as well.

Convex reformulation of 0-1 QCQP

By analogy with the original QCR method, we seek to transform a 0-1 QCQP instance into another 0-1 QCQP instance, that has the following four properties:

- the objective and constraint functions are convex
- the set of feasible solutions is identical to that of the original instance
- the cost of each feasible solution is the same as it was in the original instance...

Convex reformulation of 0-1 QCQP

By analogy with the original QCR method, we seek to transform a 0-1 QCQP instance into another 0-1 QCQP instance, that has the following four properties:

- the objective and constraint functions are convex
- the set of feasible solutions is identical to that of the original instance
- the cost of each feasible solution is the same as it was in the original instance...
- ... and the lower bound from the continuous relaxation is equal to the semidefinite/Lagrangian bound for the original instance.

0-1 QCQP with Quadratic Equations

For the case in which all constraints are equations (i.e., $N_2 = \emptyset$), we propose the following scheme:

1. Solve the SDP, obtaining optimal dual vectors $\lambda^* \in \mathbb{R}^{|N_1|}$ and $\nu^* \in \mathbb{R}^n$;
2. Perturb the objective function by adding terms of the form $x^T \mathrm{Diag}(\nu^*) x - (\nu^*)^T x$ and $\lambda_j^* (x^T Q^j x + c^j \cdot x - b_j)$ for $j \in N_1$;
3. Replace each quadratic equation with two quadratic inequalities;
4. Convexify the quadratic inequalities arbitrarily (e.g., using the minimum eigenvalue method).

0-1 QCQP with Quadratic Equations (cont.)

Theorem

When the above scheme is applied to a 0-1 QCQP instance with $N_2 = \emptyset$, the reformulated instance is convex. Moreover, if the primal SDP satisfies the Slater condition, then the lower bound from the continuous relaxation of the reformulated instance will equal the SDP bound.

Handling Quadratic Inequalities: two difficulties

The case of quadratic inequalities is more tricky because of two serious difficulties!

The first can be shown considering the following simple instance:

$$\begin{array}{lll} \max & x_1 + \cdots + x_n & \\ \text{s.t.} & x_i x_j \leq 0 & (1 \leq i < j \leq n) \\ & x_i^2 - x_i = 0 & (i = 1, \ldots, n). \end{array}$$

The optimum is 1 and one can check that the SDP bound is also 1.

Handling Quadratic Inequalities: two difficulties (cont.)

The quadratic inequalities are non-convex, so we must use the equations $x_i^2 - x_i = 0$ to convexify them.

It turns out that the best we can do is to replace the inequalities with:

$$\frac{1}{2}(x_i^2 - x_i) + \frac{1}{2}(x_j^2 - x_j) + x_i x_j \leq 0 \qquad (1 \leq i < j \leq n).$$

But the optimal solution to the continuous relaxation of the reformulated instance is then $(1/2, \ldots, 1/2)^T$, yielding a very poor upper bound of $n/2$.

Handling Quadratic Inequalities: two difficulties (cont.)

In other words, when quadratic inequalities are present:

$$\begin{aligned} \text{opt} &\geq \text{SDP bound} = \text{dual SDP bound} = \text{Lagrangian bound} \\ &\geq \text{LP bound from reformulated instance} \end{aligned}$$

and the last inequality is usually strict!

An intuitive explanation for this phenomenon is that quadratic inequalities cannot be used to perturb the objective function.

Handling Quadratic Inequalities: two difficulties (cont.)

The second difficulty is that finding an optimal reformulation is itself a rather complex optimisation problem. It is possible to formulate the problem of finding the best such reformulation as one huge SDP. Unfortunaley, the size and complexity of this SDP grows rapidly as the number of inequalities increases.

Handling Quadratic Inequalities: possible remedies

If we are willing to convert the 0-1 QCQP into a mixed 0-1 QCQP, then we can do better.

We can add continuous slack variables to convert every quadratic inequality into a quadratic equation.

Once this is done, it is always possible to reformulate the mixed 0-1 QCQP instance in such a way that the bound from the continuous relaxation is equal to the Lagrangian bound.

As before, if the primal SDP satisfies the Slater condition, this bound will be equal to the SDP bound.

Handling Quadratic Inequalities: possible remedies (cont.)

The above scheme has one disadvantage: we have to add $|N_2|$ continuous variables.

This is unattractive when $|N_2|$ is large.

We have another variant in which only one continuous variable needs to be added.

(The basic idea is to use the dual SDP solution to construct a 'surrogate' quadratic inequality that contains all of the information we need to get a good bound, and then assign a slack variable only to the surrogate constraint.)

General-Integer Variables

Next, we move on to the case of general-integer variables.

A first observation is that integer variables present no problem for the semidefinite and Lagrangian relaxations.

As for reformulation, note that we cannot use terms of the form $x_i^2 - x_i$ to perturb the objective or constraints coefficient of integer variables. This can prevent the existence of a convex reformulation.

General-Integer Variables (cont.)

To get around this, we can use binary expansion.

In fact, if such a variable is nonnegative and bounded from above by 2^p for some positive integer p, then it can be replaced by p binary variables.

Then, if all variables are bounded in this way, we can reduce the problem to 0-1 QCQP.

We show that this transformation cannot make the SDP bound any worse, and in fact in some cases it can make it better!

General-Integer Variables (cont.)

Consider the following instance with just one integer variable:

$$\min \left\{ x_1^2 - 3x_1 : \; x_1 \leq 3, \; x_1 \in \mathbb{Z}_+ \right\}.$$

There are two optimal solutions of cost -2, obtained by setting x_1 to either 1 or 2. The basic SDP relaxation is:

$$\min \left\{ X_{11} - 3x_1 : \; 0 \leq x_1 \leq 3, \; Y \succeq 0 \right\}.$$

The optimal solution to this relaxation is to set x_1 to 1.5 and X_{11} to 2.25, giving a lower bound of -2.25.

General-Integer Variables (cont.)

Now, applying binary expansion to this instance, we use the substitution $x_1 = \tilde{x}_1^1 + 2\tilde{x}_1^2$, where $\tilde{x}_1^1$ and $\tilde{x}_1^2$ are new binary variables. The SDP relaxation for the transformed instance is:

$$
\begin{array}{ll}
\min & \tilde{X}_{11}^{1,1} + 4\tilde{X}_{11}^{2,2} + 4\tilde{X}_{11}^{1,2} - 3\tilde{x}_1^1 - 6\tilde{x}_1^2 \\
\text{s.t.} & \tilde{x}_1^1 = \tilde{X}_{11}^{1,1} \\
 & \tilde{x}_1^2 = \tilde{X}_{11}^{2,2} \\
 & \tilde{Y} \succeq 0.
\end{array}
$$

One can check with an SDP solver that all optimal solutions to this SDP have cost -2. So, the transformed SDP yields an improved lower bound.

General-Integer Variables (cont.)

If, on the other hand, the general-integer variables are not bounded from above a priori, then we run into a serious problem:

Theorem (Jeroslow 1973)

Testing feasibility of an integer program with quadratic constraints is undecidable.

So we cannot expect an efficient convexification procedure in general.

(Of course, in practical applications, variables can usually be bounded quite easily.)

Continuous Variables

Finally, we consider the case in which continuous variables are permitted.

In the context of the semidefinite and Lagrangian approaches, continuous variables can be handled in exactly the same way as integer variables, and therefore they present no problems.

When it comes to reformulation, on the other hand, continuous variables cause serious difficulties. (Indeed, a convex reformulation may not exist in general, and we cannot get around this issue by *binarisation*, as we do for integer variables.)

Continuous Variables (cont.)

Here, three outcomes are possible:

- The instance can be convexified, and the bound from the continuous relaxation is as good as the SDP bound.
- The instance can be convexified, but the bound from the continuous relaxation is worse than the SDP bound.
- The instance cannot be convexified at all.

We have some necessary and sufficient conditions for these three cases to hold.

Conclusions

Basically, we have three possible situations:

- When there are no quadratic inequalities and no continuous variables, and all general-integer variables are bounded, we can always obtain a convex reformulation that gives a strong bound.
- When quadratic inequalities are present, this may not be possible, unless we are prepared to add slack variables.
- When continuous variables are present, a convex reformulation is not even guaranteed to exist.