5
A New Bounding Method

The main idea of the new method came to us when we were trying to figure out what can be done if we knew an upper bound for the optimum value of UBQP.

The method itself is explained in the next sections.

5.1 "Direct computation"

If we know an upper bound U, we can solve the following problem:

$$\begin{aligned} \min \quad & z = c^T x \\ \text{s.t.:} \quad & \frac{1}{2}x^T Q x - b^T x \leq U \end{aligned} \tag{5.1}$$

Which is "minimize a linear function subject to a quadratic constraint".

Since we know that $\min \frac{1}{2}x^T Q x - b^T x$ is a lower bound to the original problem, we know that z is a lower bound for the value of $c^T x^*$, where x^* is the solution for the UBQP.

If we know how to solve it, there is a number of interesting results we can find from it using different vectors c. Some of them are:

1. If $c = (0, \ldots, 0, 1, 0, \ldots, 0) = e_i$ and $z > 0$, then we can fix $x_i = 1$, since for all solutions whose value is less then U (and hence, for the optimal solution), $x_i > 0$.
2. If $c = (0, \ldots, 0, -1, 0, \ldots, 0) = -e_i$ and $z > -1$ we can fix $x_i = 0$.
3. With $c = (1, \ldots, 1, \ldots, 1)$, then z is a lower bound for the size of the set $\{i | x_i = 1\}$.
4. With $c = (-1, \ldots, -1, \ldots, -1)$, then $n + z$ is a lower bound for the size of the set $\{i | x_i = 0\}$

It is easy to use this information in a branch-and-bound environment, for pruning reasons.

Of course we can use the same idea of section 3.3 and solve it for $Q' = Q + 2Diag(\lambda)$ and $b' = b + \lambda$ for some λ to get even better results.

In section 5.1.1 we will show how to actually solve (5.1) if Q is positive definite.

5.1.1
Minimizing a linear function subject to a convex quadratic constraint

The following theorem shows how to optimize (5.1) in the special case that Q is positive definite. Since we will use the same idea as in section 3.3, we may assume it.

Theorem 12 *It is possible to minimize a linear function subject to one convex quadratic constraint in polynomial time.*

Proof. Let P be the problem stated at (5.1), with Q symmetric positive definite.

The function $\frac{1}{2}x^TQx - b^Tx$ is convex, so the feasible region is compact.

Since the feasible region is compact and the objective function is continuous the optimum exists. Since the objective function is linear, the optimum is reached on the boundary of the feasible region.

By the Fritz John optimality condition, there must exist $\mu_1, \mu_2 \in R$ such that $\mu_1 c^T x^* = \min(\mu_1 c^T x + \mu_2(\frac{1}{2}x^TQx - b^Tx - U))$, where x^* is the optimal solution.

Since $\mu_2 \neq 0$ (if $\mu_2 = 0$, the problem is unbounded. We already know this is not the case), we can set $\mu = \frac{\mu_1}{\mu_2}$, and $\mu c^T x^* = \min(\mu c^T x + (\frac{1}{2}x^TQx - b^Tx - U))$.

So, to solve $\min \mu c^T x + (\frac{1}{2}x^TQx - b^Tx - U)$ we can differentiate at x arriving at: $\mu c + (Qx^* - b) = 0$.

$$\mu c + (Qx^* - b) = 0 \tag{5.2}$$
$$Qx^* = b - \mu c \tag{5.3}$$
$$x^* = Q^{-1}(b - \mu c) \tag{5.4}$$

As seen in (5.4), there is a whole line that satisfies the Fritz John condition. It rest to determine the intersection of this line with the boundary of the feasible region.

$$
\begin{aligned}
\frac{1}{2}x^{*T}Qx^* - b^T x^* &= U && (5.5)\\
\frac{1}{2}(Q^{-1}(b-\mu c))^T QQ^{-1}(b-\mu c) - b^T Q^{-1}(b-\mu c) &= U && (5.6)\\
\frac{1}{2}\mu^2 c^T Q^{-1}c - \frac{1}{2}b^T Q^{-1}b &= U && (5.7)\\
\mu^2 c^T Q^{-1}c &= 2U + b^T Q^{-1}b && (5.8)\\
\mu &= \pm\sqrt{\frac{2U + b^T Q^{-1}b}{c^T Q^{-1}c}} && (5.9)
\end{aligned}
$$

So we may first calculate μ by (5.9) and then use it on (5.4) to find the optimum in polynomial time.

It should be noted that there are two different μ's, one corresponding to the minimum and one to the maximum.

Since the optimum value for μ is $c^T Q^{-1}b - \mu c^T Q^{-1}c$ and Q^{-1} is positive definite, the positive μ will give the desired minimum.

■

If we know the Cholesky factorization of Q it is easy to optimize 5.1 in $O(n^2)$ by calculating $Q^{-1}c$ and $Q^{-1}b$.

5.2 "Inverse computation"

In section 5.1 the problem of minimizing a linear function subject to a convex quadratic constraint was considered. In terms of UBQP, this means: "What is the minimum value of $c^T x^*$ given an upper bound K?".

We can now consider a somewhat different problem: "What is the best (lowest) upper bound U' that we can know, such that $c^T x$ can be less than or equal to l?", that means that if the actual known upper bound is better (less) than U', $c^T x$ must be greater than l.

First let's think about the utility of such problem. Suppose we can indeed find U'. What can we do with it?

To partially answer this question, let me show some special pairs (c, l):

1. $c = (0, \ldots, 0, 1, 0, \ldots, 0) = e_i$, $l = 0$. The solution (U') for this problem will be called $U_0(i)$.
2. $c = (0, \ldots, 0, -1, 0, \ldots, 0) = -e_i$, $l = -1$. The solution for this problem will be called $U_1(i)$.

We can find all $U_0(i)$ and $U_1(i)$ by computing the Cholesky Decomposition of a $n \times n$ matrix followed by a $O(n^3)$ computation (fast in practice). The details are explained in section 5.2.1.

Suppose we know all $U_0(i)$ and $U_1(i)$, and also an upper bound U. Now we can:

1. fix variables, achieving the same result as in section 5.1.
   If $U_0(i) > U$ then $x_i = 1$ and if $U_1(i) > U$ then $x_i = 0$.
2. get lower bounds. For all i, $\min(U_0(i), U_1(i))$ is a lower bound.
   So $\max_i(\min(U_0(i), U_1(i)))$ is also a lower bound for the problem.
3. decide on the branching variable (more details in section 6.2).
4. guide heuristics (see section 6.4).

5.2.1
Solving the "inverse problem"

Theorem 13 *It is possible to solve the problem stated at 5.2*

Proof. Let P be the problem stated at (5.1), with Q symmetric positive definite.

The feasible set shrinks continuously as U gets smaller, and also, the optimal value gets continuously greater. So the "largest U such that $c^T x$ can be less or equal to l" is the unique U for which the optimal value of P is l.

Using the results from 5.1.1, it follows that:

$$c^T Q^{-1} b - \mu c^T Q^{-1} c = l \qquad \text{From (5.4)} \tag{5.10}$$

$$\mu^2 = \left(\frac{c^T Q^{-1} b - l}{c^T Q^{-1} c} \right)^2 \tag{5.11}$$

$$\mu^2 = \frac{2U + b^T Q^{-1} b}{c^T Q^{-1} c} \qquad \text{From (5.9)} \tag{5.12}$$

$$\frac{2U + b^T Q^{-1} b}{c^T Q^{-1} c} = \left(\frac{c^T Q^{-1} b - l}{c^T Q^{-1} c} \right)^2 \qquad \text{From (5.11) and (5.12)} \tag{5.13}$$

$$2U + b^T Q^{-1} b = \frac{\left(c^T Q^{-1} b - l\right)^2}{c^T Q^{-1} c} \tag{5.14}$$

$$U = \frac{1}{2} \left(\frac{\left(c^T Q^{-1} b - l\right)^2}{c^T Q^{-1} c} - b^T Q^{-1} b \right) \tag{5.15}$$

∎

Again, if we know the Cholesky factorization of Q it is easy to solve the "inverse problem" in $O(n^2)$ by calculating $Q^{-1}c$ and $Q^{-1}b$.

Of course, the same idea used on 3.3 can be used here, and we can solve the problem for several values of λ.

5.3 Further Improving the Bound

We can further improve the results of sections 5.1 and 5.2 by adding the constraints $0 \leq x_i \leq 1$.

Get, for instance, the problem in 5.1 and augment it with these constraints.

$$
\begin{aligned}
\min \quad & z = c^T x && && (5.16)\\
\text{s.t.:} \quad & \frac{1}{2} x^T Q x - b^T x \leq U && && \\
& -x_i \leq 0 && \forall i \in \{0, \ldots, n\} && (5.17)\\
& x_i - 1 \leq 0 && \forall i \in \{0, \ldots, n\} && (5.18)
\end{aligned}
$$

We can get the Lagrangian dual of the above problem, with respect to (5.17) and (5.18), and arrive to the following:

$$
\begin{aligned}
\mathcal{L}_D(\alpha, \beta) \quad & = \min c^T x - \alpha^T(-x) - \beta^T(x - 1)\\
& = \min (c + \alpha - \beta)^T x + \sum_i \beta_i\\
\text{s.t.:} \quad & \frac{1}{2} x^T Q x - b^T x \leq U
\end{aligned}
$$

So, one can compute $\mathcal{L}_D(\alpha, \beta)$ for a given $(\alpha, \beta) \in \mathbb{R}^{2n}$ in exactly the same way described in 5.1.1 (by letting $c = (c + \alpha - \beta)$). Also, the dual function is always concave [ruszczynski2006].

Then, the dual problem:

$$
\begin{aligned}
\max \quad & \mathcal{L}_D(\alpha, \beta) \\
\text{s.t.:} \quad & \alpha \geq 0 \\
& \beta \geq 0
\end{aligned}
$$

can be solved by a convex optimization method. Note that it is easy to differentiate $\mathcal{L}_D$.

This dual problem will give a lower bound on the value of (5.16), which can be used to the same purpose as the lower bounds on (5.1).

A similar approach can be used to improve the problem of section 5.2.