Large-Scale Quadratically Constrained Quadratic Program via Low-Discrepancy Sequences

Kinjal Basu, Ankan Saha and Shaunak Chatterjee
LinkedIn Corporation

Problem

We consider the following quadratically constrained quadratic programming (QCQP) problem,

$$\begin{aligned} \underset{\mathbf{x}}{\text{Minimize}} \quad & (\mathbf{x}-\mathbf{a})^T\mathbf{A}(\mathbf{x}-\mathbf{a}) \\ \text{subject to} \quad & (\mathbf{x}-\mathbf{b})^T\mathbf{B}(\mathbf{x}-\mathbf{b}) \leq \tilde{b}, \\ & \mathbf{C}\mathbf{x} = \mathbf{c}. \end{aligned} \tag{1}$$

where $\mathbf{A}, \mathbf{B}$ are $n \times n$ positive-definite matrices.

The usual techniques include SDP and RLT relaxations but both convert the problem from $O(n)$ to $O(n^2)$ variables. This makes solving such problems extremely expensive in large-scale applications.

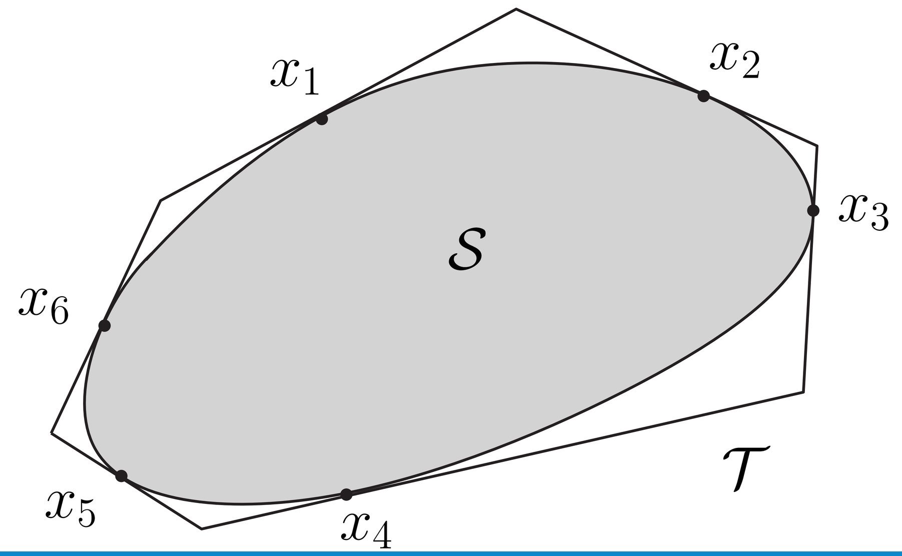
QCQP to QP Approximation

To get a scalable solution, we approximate the quadratic constraint by a set of linear constraints thus, obtaining a quadratic program (QP) with n variables.

We choose a set of N points such that each point $\mathbf{x}_j$ belongs to the boundary. The transformed problem $\mathcal{P}(\mathcal{X}_N)$ can be written as

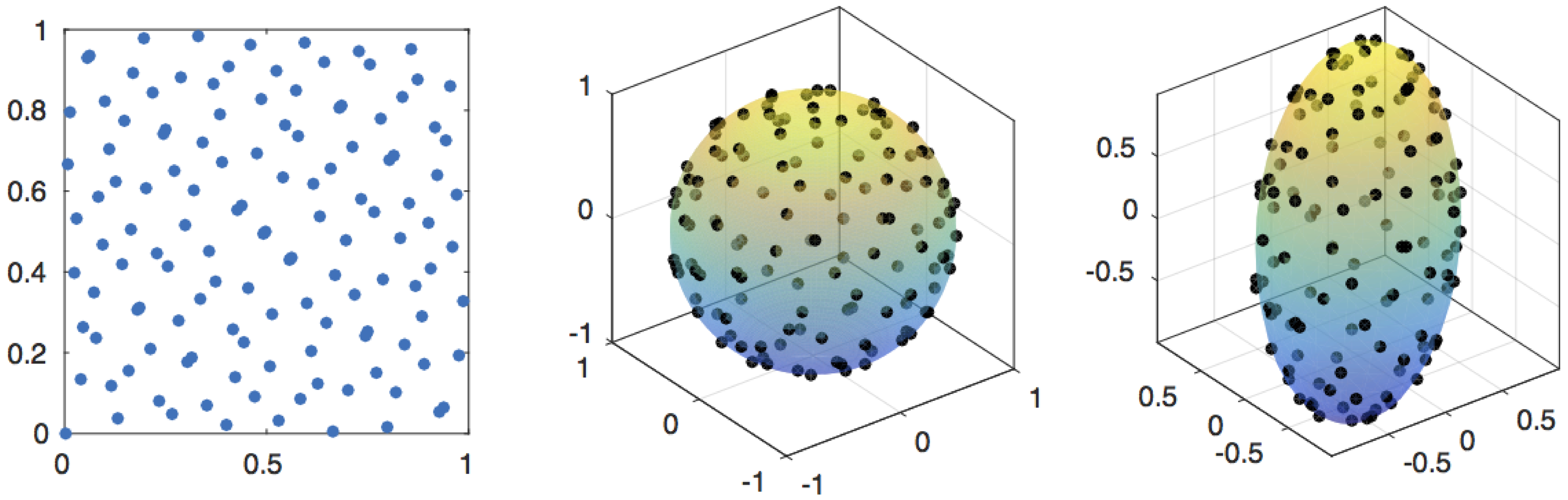
$$\begin{aligned} \underset{\mathbf{x}}{\text{Minimize}} \quad & (\mathbf{x}-\mathbf{a})^T\mathbf{A}(\mathbf{x}-\mathbf{a}) \\ \text{subject to} \quad & (\mathbf{x}-\mathbf{b})^T\mathbf{B}(\mathbf{x}_j-\mathbf{b}) \leq \tilde{b} \\ & \text{for } j = 1, \ldots, N \\ & \mathbf{C}\mathbf{x} = \mathbf{c}. \end{aligned} \tag{2}$$

For example, the tangent planes through the 6 points $\mathbf{x}_1, \ldots, \mathbf{x}_6$ create the approximation to the quadratic constraint $\mathcal{S}$ in two dimensions.



Low-Discrepancy Sampling

The accuracy of the solution to $\mathcal{P}(\mathcal{X}_N)$ is dependent on the choice of the points $\mathcal{X}_N$. Choosing random points can lead to arbitrary bad solutions. To get an accurate solution we resort to optimally mapping a low-discrepancy sequence to the ellipsoidal constraint, which has good equidistribution property. We use a (t, m, s)-net as a starting point on the unit hypercube which is then mapped to the surface of the ellipsoid via a measure preserving map.



The left panel shows a $(0, 7, 2)$-net in base 2 which is mapped to a sphere in 3 dimensions (middle panel) and then mapped to the ellipsoid as seen in the right panel.

Experimental Results

We consider random objective functions with the true global minimum outside of the constraint domain. SDP and Exact (Interior point methods) give us the true optimal. For large n the algorithms do not converge in time (1 hour). Our sampling scheme give much closer objective value to the truth than other sampling techniques.

Table 1: Optimal objective value and convergence time

n	Our method	Sampling on $[0,1]^n$	Sampling on $\mathbb{S}^n$	SDP	RLT	Exact
5	3.00 (4.61s)	2.99 (4.74s)	2.95 (6. 11s)	**3.07** (0.52s)	3.08 (0.51s)	**3.07** (0.49)
50	99668 (15.55s)	15122 (18.98s)	26239 (17.32s)	$\mathbf{1.11 \times 10^5}$ (4.31s)	1.08×10^5 (2.96s)	$\mathbf{1.11 \times 10^5}$ (0.64)
100	1.40×10^6 (58.41s)	69746 (1.03m)	1.24×10^6 (54.69s)	$\mathbf{1.62 \times 10^6}$ (30.41s)	1.52×10^6 (15.36s)	$\mathbf{1.62 \times 10^6}$ (2.30s)
10^5	3.10×10^8 (25.82m)	7.12×10^7 (24.59m)	8.39×10^7 (27.23m)	NA	NA	NA
10^6	3.91×10^9 (38.30m)	2.69×10^8 (39.15m)	7.53×10^8 (37.21m)	NA	NA	NA

Convergence Results

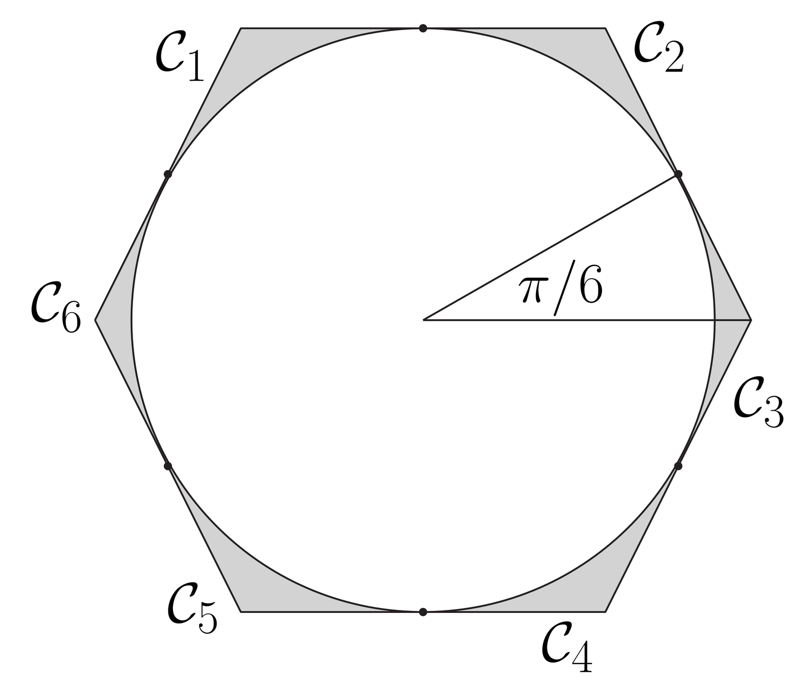
Theorem 1 $\lim_{N\to\infty} \|\mathbf{x}^*(N) - \mathbf{x}^*\| = 0$

Theorem 2 *If $\|\mathbf{x}^*(N) - \mathbf{x}^*\| = O(g(N))$, then $|f(\mathbf{x}^*(N)) - f(\mathbf{x}^*)| \leq Cg(N)$ where $C > 0$ is a constant.*

For example, if $\mathcal{S}$ was the unit circle, then we have,

$$\begin{aligned} g(N) &:= \max_{i=1,\ldots,N} \sup_{\mathbf{t},\mathbf{x}:\mathcal{A}(\mathbf{t},\mathbf{x})\in\mathcal{C}_i} \|\mathbf{t}-\mathbf{x}\| \\ &= \tan\left(\frac{\pi}{N}\right) = O\left(\frac{1}{N}\right). \end{aligned} \tag{3}$$

Combining this observation with Theorem 2 shows that in order to get an objective value within ϵ of the true optimal, we would need N to be a constant multiplier of ϵ^{-1}.



Six equivalent conic regions for a unit circle.

Future Work

- Comparison with commercial solvers such as CPLEX and large-scale SDP solvers based on ADMM such as Splitting Conic Solver (SCS)
- Finding explicit bounds for common domains.
- Finding accurate rates by considering the growth of the eigenvalues of the matrices.