Convexification with bounded gap for randomly projected QO and beyond.

T. Fuji[1], P. L. Poirion[2], A. Takeda[3]

[1] Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan
teru0818258@g.ecc.u-tokyo.ac.jp

[2] Center for Advanced Intelligence Project, RIKEN, Tokyo 103-0027, Japan.
pierre-louis.poirion@riken.jp

[3] Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan and Center for Advanced Intelligence Project, RIKEN, Tokyo 103-0027, Japan
takeda@mist.i.u-tokyo.ac.jp

Mots-clés : *random projection, Johnson-Lindenstrauss lemma, quadratic programming, non-convex optimization*

We consider the following non-convex quadratic optimization problem :

$$\mathcal{P} \equiv \min_{x} \{x^\top Q x + c^\top x \mid Ax \leq b\}, \tag{1}$$

where $Q \in \mathbb{R}^{n \times n}$ is a symmetric matrix, $A \in \mathbb{R}^{m \times n}, c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. Here, we assume that the feasible region is full dimensional (hence, equality constraints are excluded) and at least one of the eigenvalues of Q is positive. Here we are interested in the case where the dimension n of the variable x is large.

In this talk, we show that we can find a feasible solution with a bounded approximation error to the optimal value of (1). To do so, we reduce the non-convex problem to a lower-dimensional convex optimization problem using random projection and convexification techniques, and evaluate the gap between optimum values of the two optimization problems. Random projection refers to the technique that maps a set of points $X \subsetneq \mathbb{R}^n$ to a set $PX \subsetneq \mathbb{R}^d$ in a lower dimensional subspace with random matrices $P \in \mathbb{R}^{d \times n}$ in a way that some intrinsic properties of the set X are approximately preserved with high probability. The main idea of random projections comes from the Johnson-Lindenstrauss lemma [2] that states that if the probability distribution of P is properly chosen then there exists $d < n$ such that the Euclidean distance between any pair $x, y \in X$ is approximately preserved with high probability, i.e. $\|Px - Py\| \approx \|x - y\|$.

Here we will use random projection to define a convexification of (2) and give some error bounds for the error between these two problems. Notice that in [1] the authors already use a random projection matrix $P \in \mathbb{R}^{d \times n}$ to project (1) into the following QP :

$$\mathcal{RP} \equiv \min_{u} \{u^\top \bar{Q} u + \bar{c}^\top u \mid \bar{A} u \leq b\}, \tag{2}$$

where $u \in \mathbb{R}^d$, $\bar{Q} = PQP^\top, \bar{c} = Pc$ and $\bar{A} = AP^\top$. However, although (2) is a QP of smaller size, i.e. the variables of (2) belong to a smaller dimensional space, if (1) is non-convex then the projected problem will also almost surely be non-convex, and hence will be hard to solve. Therefore, we focus on the fact that if d is small enough then eigenvalues of the matrix $\bar{Q}$ are skewed towards positive values, which implies that ignoring negative eigenvalues for the reduced matrix due to the convexification does not lose much information about problem (1). In this paper, taking advantage of this fact, we show the following : if the dimension d is carefully chosen then (1) can be approximated by a convex QP of smaller size. More precisely, we consider the following convex QP :

$$\mathcal{CRP} \equiv \min_{u}\{u^\top \bar{Q}^+ u + \bar{c}^\top u \mid \bar{A}u \leq b\}, \tag{3}$$

where $\bar{Q}^+ = F^+(\bar{Q})$ is the projection of $\bar{Q}$ onto the positive semidefinite cone. Using an optimum u^* of (3), we have a feasible solution $P^\top u^*$ to (1), for which an approximation error from the optimum value of (1) is estimated.

To the best of our knowledge, this is the first work to use random projection for convexification of non-convex optimization problems. The plan of the talk is the following :

We will begin by introducing mathematical preliminaries of random projections, then, we prove our main results on approximate optimality under the assumption that $trace(Q) > 0$, then we discuss how to relax the assumption while achieving similar theoretical results. Finally, we discuss the results of numerical experiments for two types of problems : randomly generated problems and support vector machine (SVM) problems with indefinite kernel which are attributed to non-convex quadratic optimization problems.

Références

[1] Claudia D'Ambrosio, Leo Liberti, Pierre-Louis Poirion, and Ky Vu. Random projections for quadratic programs. *Math.Program.*, 183 :619–647, 2020.

[2] W. Johnson and J. Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In G. Hedlund, editor, *Conference in Modern Analysis and Probability*, volume 26 of *Contemporary Mathematics*, pages 189–206, Providence, 1984. American Mathematical Society.