ROADEF 2020, Random projections for Linear Programming with inequalities

Pierre-Louis Poirion[1], Bruno F. Lourenço[2], Akiko Takeda[1,2]

[1] RIKEN Center for Advanced Intelligence Project, Tokyo, Japan
pierre-louis.poirion@riken.co.jp
[2] University of Tokyo, Tokyo, Japan
{lourenco,takeda}@mist.i.u-tokyo.ac.jp

Mots-clés : *Projections aléatoires, optimisation linéaire.*

1 Introduction

In this paper we propose a new random projection method that allows to reduce the number of inequalities of a Linear Program (LP). More precisely, we randomly aggregate the constraints of a LP into a new one with much fewer constraints, while preserving, approximatively, the value of the LP. This extends the work in [4], where the authors considered an LP with equality constraints and proved that we can build a new LP, whose value approximate the original one, with much fewer constraints by randomly aggregating them using a matrix of independent and identically distributed (iid) random variables. The results obtained where derived from the Johnson-Lindenstrauss Lemma [1, 2], which states that a set of high-dimensional points can be projected to a much lower dimensional one, while preserving, approximatively, the Euclidean distance between these points. The extension, to the inequality case, proposed in this paper is non-trivial as we need to consider random matrices with non-negative entries.

2 Random matrices applied to LP

Random matrices are matrices $T \in \mathbb{R}^{k\times m}$ whose entries a drawn from a probability distribution. When the underlying distribution is properly chosen, these matrices can have some very interesting properties : the Johnson-Lindenstrauss Lemma (JLL), [1, 2], states that, if the entries of T are drawn independently from the standard normal distribution, $\mathcal{N}(0, \frac{1}{k})$, it is possible to project a set of n points of $\mathbb{R}^m$ into a space of dimension $k = O(\frac{\log(n)}{\varepsilon^2})$ while preserving approximately (with "ε precision"), with arbitrarily high probability (w.a.h.p.), the Euclidean distance between these points.
Recently, this result has been exploited, [4], to prove that equality constraints of an LP written in standard form, could be randomly aggregated, using a random matrix T with $k < m$, into a new LP :

$$\begin{cases} \min\limits_x & c^\top x \\ & Ax = b \\ & x \in \mathbb{R}^n_+ \end{cases} \qquad\qquad \begin{cases} \min\limits_x & c^\top x \\ & TAx = Tb \\ & x \in \mathbb{R}^n_+ \end{cases}$$

while preserving, approximately, w.a.h.p. the optimal value of the problem. Considering the dual setting, this result allows to reduce the dimension of the dual problem where the dual variables $y \in \mathbb{R}^m$ are "replaced" by $T^\top y_T$ with $y_T \in \mathbb{R}^k$. This framework however, does not allow to reduce both the number of constraints and the number of variables of an LP as in the dual problem we now have inequality constraints instead of equality ones. Indeed, to randomly aggregate a set of inequality constraints : $Ax \leq b$, we need a random matrix S whose entries

S_{ij} are non-negative.

In this paper, we propose the first method that allows to randomly aggregate a set of inequality constraints in an LP. More precisely, let us consider the pair :

$$\mathcal{P}\left\{\begin{array}{ll}\min\limits_x & c^\top x\\ & Ax \geq b\\ & x\in\mathbb{R}^n\end{array}\right. \quad (1) \qquad \mathcal{P}_S\left\{\begin{array}{ll}\min\limits_x & c^\top x\\ & SAx \geq Sb\\ & x\in\mathbb{R}^n\end{array}\right. \quad (2)$$

with $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^n$ and where $S \in \mathbb{R}^{k\times m}$ is a random *iid* matrix such that $S_{ij} = T_{ij}^2$ where T_{ij} is drawn from the normal distribution $\mathcal{N}\left(0, \frac{1}{k}\right)$. Although this looks very similar to the equality case it is actually quite different : indeed the random matrix S does not satisfy the JLL property, hence a different analysis should be applied. Notice that since each entry of S is non-negative, $\mathcal{P}_S$ is a relaxation of $\mathcal{P}$, hence $v(\mathcal{P}_S) \leq v(\mathcal{P})$ (where $v(\cdot)$ denotes the optimal value of an optimization problem). The difficult part is to prove that there exists a decreasing function $\delta(\varepsilon) > 0$ (recall that $k = O(\frac{\log(n)}{\varepsilon^2})$) such that, w.h.a.p.,

$$v(\mathcal{P}) - \delta(\varepsilon) \leq v(\mathcal{P}_S) \leq v(\mathcal{P}).$$

In this talk we will prove that, under an additional assumption, the results obtained in [4] can be extended to the inequality case. We will also present some numerical results that support the theory.

Références

[1] S. Dasgupta and A. Gupta. An elementary proof of a theorem by Johnson and Lindenstrauss. *Random Structures and Algorithms*, 22 :60–65, 2002.

[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.

[3] Roman Vershynin. High-dimensional probability, an introduction with applications in data science. 2018.

[4] Ky Khac Vu, Pierre-Louis Poirion, and Leo Liberti. Random projections for linear programming. *Math. Oper. Res.*, 43(4) :1051–1071, 2018.