Are Random Projections really useful in Mathematical Programming ?

Leo Liberti[1]

LIX CNRS Ecole Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau, France
leo.liberti@cnrs.fr

Mots-clés : *dimensional reduction, Johnson-Lindenstrauss Lemma, large-scale*

The usefulness of Mathematical Programming (MP), as a formal language for describing optimization problems, resides in its solvers. Instead of devising a new algorithm for each new problem, just model it using MP then deploy a solver on it. The smaller the class targeted by the solver, and, as long as it contains the problem at hand, the higher the chances of obtaining a good or even certified optimal solution. This process is now streamlined to a point that there are high-level languages for MP (e.g. AMPL [4]), that automatically interface with most high-quality solvers (e.g. [5, 2, 10]).

The issue arises when the problem instances exceed a certain size. Then all bets are off, solvers can at most be used as heuristics, if at all. If for structured Linear Programs (LP) the threshold size is in the millions of variables, dense unstructured LPs and most other MP subclasses have much lower thresholds : from $O(10^4)$ for some conic programs to $O(100)$ or even $O(10)$ for complicated Mixed-Integer Nonlinear Programs (MINLP).

This is where Random Projections (RP) come in. RPs are simply $k \times m$ random matrices T, sampled componentwise from subgaussian distributions, that decrease the dimensionality of a vector set $X \subset \mathbb{R}^m$ from m to $k \ll m$. The best known result in this area is the Johnson-Lindenstrauss Lemma (JLL) [6], which applies to finite sets X with $|X| = n$, and states that for a given $\epsilon \in (0, 1)$, if $k = O(\epsilon^{-2} \ln n)$, then TX is ϵ-approximately congruent to X with high probability, i.e.

$$\forall i < j \leq n \quad (1 - \epsilon)\|X_i - X_j\|_2 \leq \|TX_i - TX_j\|_2 \leq (1 + \epsilon)\|X_i - X_j\|_2$$

holds with probability exceeding $1 - n(n-1)e^{\mathcal{C}\epsilon^2 k}$, where $\mathcal{C}$ is a “universal constant”.

The application of RPs to vectors is now well understood [11]. Their applications to optimization problems with objective functions involving the ℓ_2 norm and convex constraints is reasonably well understood [1, 9]. Other problems often occurring in Machine Learning (ML) can also be treated with RP. The application of RPs to whole subclasses of MP, on the other hand, is recent : in past works, we addressed LP [13], Semidefinite Programming (SDP) [8], Quadratic Programs (QP) [12, 3], and some problems involving integer variables [14, 7].

The application of RPs to MP poses two serious issues, which require a detailed treatment for each separate MP subclass : the fact that MP formulations are symbolic rather than numerical entities, and the fact that a decision variable vector embodies a potentially uncountable set of vectors : this makes the application of the JLL impossible, since it requires X to be finite.

Besides the theoretical advances we proposed in past works, most of our papers also carry substantial computational sections, where we show to what extent one can hope to apply RPs to a MP formulation and obtain a useful result.

My talk will focus on the empirical side of RPs for MP. I will discuss success and failure stories, as well as technical difficulties to be overcome.

Références

[1] C. Boutsidis and P. Drineas. Random projections for the nonnegative least-squares problem. *Linear Algebra and its Applications*, 431 :760–771, 2009.

[2] COIN-OR. *Introduction to IPOPT : A tutorial for downloading, installing, and using IPOPT*, 2006.

[3] C. D'Ambrosio, L. Liberti, P.-L. Poirion, and K. Vu. Random projections for quadratic programs. *Mathematical Programming B*, 183 :619–647, 2020.

[4] R. Fourer and D. Gay. *The AMPL Book.* Duxbury Press, Pacific Grove, 2002.

[5] IBM. *ILOG CPLEX 12.10 User's Manual.* IBM, 2020.

[6] 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, RI, 1984. AMS.

[7] L. Liberti and B. Manca. Side-constrained minimum sum-of-squares clustering : Mathematical Programming and random projections. *Journal of Global Optimization*, accepted.

[8] L. Liberti, P.-L. Poirion, and K. Vu. Random projections for conic programs. *Linear Algebra and its Applications*, 626 :204–220, 2021.

[9] M. Pilanci and M. Wainwright. Randomized sketches of convex programs with sharp guarantees. In *International Symposium on Information Theory (ISIT)*, pages 921–925, Piscataway, 2014. IEEE.

[10] N.V. Sahinidis and M. Tawarmalani. *BARON 7.2.5 : Global Optimization of Mixed-Integer Nonlinear Programs,* User's Manual, 2005.

[11] S. Vempala. *The Random Projection Method.* Number 65 in DIMACS Series in Discrete Mathematics and Theoretical Computer Science. AMS, Providence, RI, 2004.

[12] K. Vu, P.-L. Poirion, C. D'Ambrosio, and L. Liberti. Random projections for quadratic programs over a Euclidean ball. In A. Lodi and *et al.*, editors, *Integer Programming and Combinatorial Optimization (IPCO)*, volume 11480 of *LNCS*, pages 442–452, New York, 2019. Springer.

[13] K. Vu, P.-L. Poirion, and L. Liberti. Random projections for linear programming. *Mathematics of Operations Research*, 43(4) :1051–1071, 2018.

[14] K. Vu, P.-L. Poirion, and L. Liberti. Gaussian random projections for Euclidean membership problems. *Discrete Applied Mathematics*, 253 :93–102, 2019.