CS 226 / 6.889 Sketching Algorithms for Big Data
— Fall 2017

Problem Set 3

Due: 11:59pm, November 8, 2017.
Submit to: `sketchingbigdata-f17-assignments@seas.harvard.edu`

See homework policy at `http://www.sketchingbigdata.org/fall17/syllabus/`

All problems 10 points each, except problem 5.

Problem 1: Michael's counterexample   In Lecture 13 we have seen that (ck, δ)-RIP implies null-space property with a constant $C = C(c, \delta) > 1$ that can be set arbitrarily close to 1 by selecting large enough c and small enough δ. In Lecture 15 we have seen that appropriately scaled unbalanced expander matrices satisfy RIP1 as well as null-space property, also with C arbitrarily close to 1. Recall that we need $C < 2$ for the L1 minimization to work.

It would be nice if we could show that (ck, δ)-RIP1, for appropriately large c and small δ, implies null space property for C close to 1. However, this is not true, and the goal of this problem is to demonstrate this. In particular, for any pair of constants $c > 1$ and $\delta > 0$, give an $m \times n$ matrix A which satisfies (c, δ)-RIP1, but not the null space property with $C < 2$.

Problem 2: Parsimonious sampling   In Lecture 16 you have seen how to recover a 1-sparse Fourier signal from two signal samples.

(a) Show that it is not possible to uniquely recover a 1-sparse Fourier signal from only a single sample (deterministically defined in advance).

(b) Show that it is not possible to uniquely recover a 2-sparse Fourier signal from only two samples (again, deterministically defined in advance).

Problem 3: Holographic (non)-sparsity   Let F be an $n \times n$ Fourier matrix. Clearly, if $\|x\|_0 = 1$, then Fx is equal to one of the columns of F multiplied by a non-zero scalar, and therefore $\|Fx\|_0 = n$. Show the following generalization: if $\|x\|_0 = k$, then $\|Fx\|_0 \geq n/k$.

Problem 4: Better L_p norm estimation   For any $p > 2$ show that if there is a linear sketching algorithm that solves L_2 sampling for vectors of dimension n using space $\log(n)^{O(1)}$, then there is a linear sketching algorithm that provides a constant factor estimation of the L_p norm for vectors of the same dimension in space $n^{1-2/p}\log(n)^{O(1)}$.

Problem 5:   (1 point) How much time did you spend on this problem set? If you can remember the breakdown, please report this per problem. (sum of time spent solving problem and typing up your solution)