On the conditioning of random subdictionaries, by Joel A. Tropp [1]

Davy Leo Gjorgjevski Martin

March 30, 2022

Many problems appearing in signal processing are in the form of a linear inverse problems. With appropriate knowledge on the signals of interest it is possible (and can be beneficial) to solve these problems under sparsity constraints. Finding solutions under sufficient sparsity constraints was proven to be well-posed and computationally tractable problem [2]. Techniques such as "Matching pursuit" and "Basis pursuit" were developed for this purpose. However, theoretical results were often constrained to pairs of bases and could only guarantee to find the solution when the sparsity constraint was of size $\mathcal{O}(\sqrt{d})$, a phenomenon known as the square-root bottleneck. On the other hand, it was known from empirical results that sparse recovery succeeded beyond the $\sqrt{d}$ range [3].

The present article shows how to relax the definition of sparsity to allow up to $\frac{d}{\log d}$ nonzero components in general dictionaries. The approach is to consider subdictionaries formed by selecting fixed number of columns from the dictionary uniformly at random. The main results are that on average, and also with large probability (under suitably chosen parameters) a random subdictionary is well conditioned. In particular, most subdictionaries have full rank. The proofs rely on tools from probability theory in Banach spaces. The key step in proving the main results is a bound on the moments of the spectral norm of a random compression of a given matrix. This bound implies that the spectral norm of the random compression is subgaussian variable and one obtains the results by a simple concentration inequality. An example where ℓ^1 minimization is used to solve (P0) is discussed. The links and differences between the results in the article and the compressed sensing theory is also discussed [4].

In order to better appreciate the results of this article, we believe that some time should be spent explaining the deterministic work on the subject. Our presentation is therefore divided into two parts: deterministic approach (results from linear algebra and convex optimization) and the probabilistic approach (techniques from probability on Banach spaces, decoupling, symmetrization and concentration inequalities).

References

[1] Joel A. Tropp. "On the conditioning of random subdictionaries". In: *Applied and Computational Harmonic Analysis* 25.1 (2008), pp. 1–24. ISSN: 1063-5203. DOI: `https://doi.org/10.1016/j.acha.2007.09.001`.

[2] David L. Donoho and Xiaoming. Huo. "Uncertainty principles and ideal atomic decomposition". In: *IEEE Transactions on Information Theory* 47.7 (2001), pp. 2845–2862. DOI: `10.1109/18.959265`.

[3] Alfred M. Bruckstein, David L. Donoho, and Michael Elad. "From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images". In: *SIAM Review* 51.1 (2009), pp. 34–81. DOI: `10.1137/060657704`.

[4] Emmanuel Candès and Justin Romberg. "Sparsity and incoherence in compressive sampling". In: *Inverse Problems* 23.3 (Apr. 2007), pp. 969–985. ISSN: 1361-6420. DOI: `10.1088/0266-5611/23/3/008`.