THE UNIVERSITY OF ARIZONA. Program in Applied Mathematics

Randomized Iterative Thresholding for Sparse Approximations

Robert Crandall[1], Ali Bilgin[2,3]

[1]Program in Applied Mathematics, [2]Department of Electrical and Computer Engineering, [3]Biomedical Engineering, The University of Arizona, AZ, USA

Introduction

- An underdetermined linear system of equations

$$y = Ax$$

has infinitely many solutions. However, it may have a unique *sparse* solution if A satisfies the Restricted Isometry Property (RIP).

- For A satisfying RIP, can use Iterative Hard Thresholding (IHT) to find x if x is sparse enough.
- If we introduce additional prior information about the system, we can derive the Minimum Mean-Squared Error (MMSE) estimator. This estimator is impractical to compute: complexity is exponential in signal length!
- Can modify IHT algorithm to approximate MMSE estimate. By intelligently combining multiple sparse solutions, we can achieve a better result!

Iterative Hard Thresholding

- IHT algorithm seeks x using iteration

$$x^{n+1} = H_k(x^n + A^T(y - Ax))$$

- H_k is the *hard thresholding* operator that sets all but the largest k elements (in magnitude) of a vector to zero
- Fast, effective method for solving noisy sparse system

$$y = Ax + v$$

when A satisfies RIP and x is sparse; guaranteed to come within $C||v||_2$ of true solution

MMSE Estimation

- Suppose we have known prior distributions p(x), p(v) and P(S) (S is support set of x). MMSE estimator is

$$x^{MMSE} = \mathrm{E}(x|y) = \sum_S P(S|y)E(x|y,S)$$

- This is a weighted sum of optimal solutions on each possible support. Too many to compute in practice; but maybe we can sample from P(S|y)!
- In paper “A Plurality of Sparse Representations is Better Than the Sparsest One Alone,” Elad & Yavneh introduce RandOMP algorithm; modified version of Orthogonal Matching Pursuit that samples randomly from E(x|y,S) and combines to approximate MMSE estimator
- RandOMP demonstrates improved performance in terms of MSE over

Randomized Thresholding

- Inspired by RandOMP, introduce randomized thresholding operator H_P for a distribution P(S). H_P zeros out all elements of x except on a random support S, chosen from P(S)
- RandIHT: same as IHT but use this randomized thresholding operator:

$$x^{n+1} = H_P(x^n + A^T(y - Ax))$$

- Goal: choose P so that RandIHT converges to a solution E(x|y,S) with probability P(S|y). This may not be feasible but we can try to approximate!

Example: Gaussian Signal and Noise

- To illustrate, suppose we know a priori that #S = k (signal is k-sparse), the nonzero entries of x are i.i.d. Gaussian, and the noise v is i.i.d. Gaussian. Define

$$Q_S = \frac{1}{\sigma_v^2} A_S^T A_S + \frac{1}{\sigma_x^2} I$$

$$E(x|y,S) = \frac{1}{\sigma_v^2} Q_S^{-1} A_S^T y$$

- Then

$$P(S|y) \propto \exp\left[\frac{1}{2} z_S^T Q_S z_S + \frac{1}{2}\log(\det(Q_S^{-1}))\right]$$

- To make a tractable algorithm: make simplifying assumption that any S columns of A are orthonormal (RIP quantifies how good this assumption is). Then

$$P(S|y) \propto \exp\left[\frac{1}{2\sigma_v^2}\frac{\sigma_x^2}{\sigma_x^2 + \sigma_v^2}\left|\left|A_S^T y\right|\right|^2\right]$$

- Choose randomized thresholding operator H_P based on this; probability of a an element being used in the support is a function of the magnitude of the correlations of y with columns of A
- Run RandIHT L times, then combine solutions by simple averaging:

$$\hat{x}^{RIHT} = \sum_{i=1}^{L} \hat{x}_i$$

Simulation Results

- For numerical simulations in Matlab we choose signal length n = 512, and measurement length m = 256. Signals are sparse with 8 non-zeros and support chosen uniformly at random. Dictionary A chosen with i.i.d. Gaussian entries, then columns are normalized in L^2.

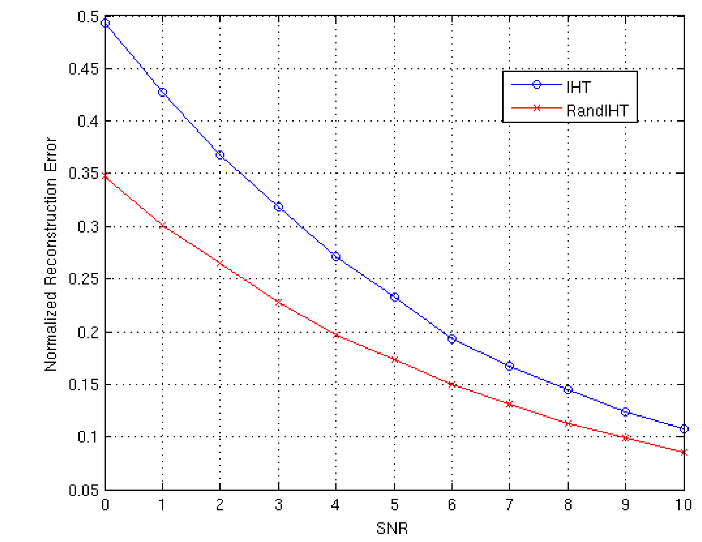

Figure 1: reconstruction error vs. SNR

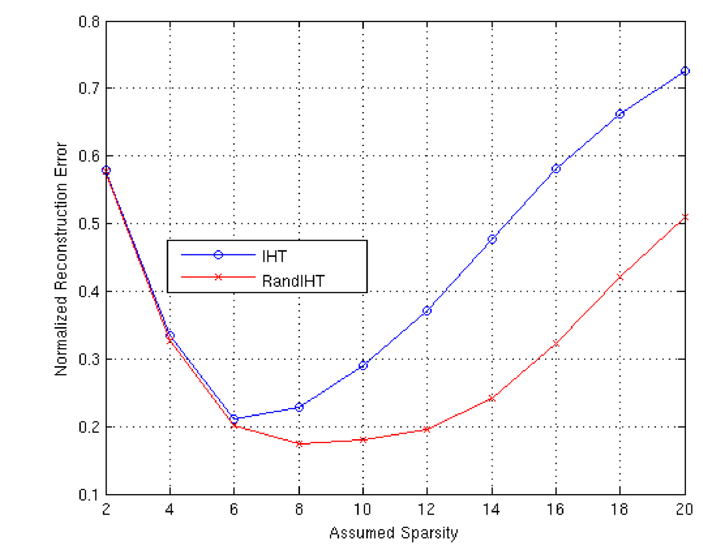

Figure 2: reconstruction error vs. assumed sparsity

- Figure 1 demonstrates improved performance across a range of SNRs. Figure 2 shows improved robustness to choice of sparsity level, which in practice may not be known
- Next we compare image reconstructions from noisy undersampled measurements using crude assumption that wavelet coefficients meet our simple Gaussian model
- Top left: IHT, 5% wavelet coefficients recovered. Top right: RandIHT, 5%. Bottom left: IHT, 15%. Bottom right: RandIHT, 15%

- Performance is similar for optimal choice of sparsity, but RandIHT is more robust to sparsity choice which is not known a priori

Summary

- Our algorithm RandIHT generates random solutions that sample from this set of candidate supports. They can be combined to approximate the MMSE estimate and provide lower mean-squared error than the IHT solution.
- Future work: convergence results? Combine with adaptive signal representations, e.g. spatially adaptive sparsity transforms (BM3D, etc)