Title:
Sparse Solution of Underdetermined Linear Equations by Stagewise Orthogonal Matching Pursuit

Author(s):
David L. Donoho, Yaakov Tsaig, Iddo Drori, and Jean-Luc Starck

Technical Report number (Dept. of Statistics, Stanford Univ.):
2006-2

Date:
April 2006

Abstract:

Finding the sparsest solution to underdetermined systems of linear equations $y = \Phi x$ is NP-hard in general. We show here that for systems with 'typical'/'random' Φ, a good approximation to the sparsest solution is obtained by applying a fixed number of standard operations from linear algebra.

Our proposal, Stagewise Orthogonal Matching Pursuit (StOMP), successively transforms the signal into a negligible residual. Starting with initial residual $r_0 = y$, at the s-th stage it forms the 'matched filter' $\Phi^T r_{s-1}$, identifies all coordinates with amplitudes exceeding a specially-chosen threshold, solves a least-squares problem using the selected coordinates, and subtracts the least-squares fit, producing a new residual. After a fixed number of stages (e.g. 10), it stops. In contrast to Orthogonal Matching Pursuit (OMP), many coefficients can enter the model at each stage in StOMP while only one enters per stage in OMP; and StOMP takes a fixed number of stages (e.g. 10), while OMP can take many (e.g. n). StOMP runs much faster than competing proposals for sparse solutions, such as ℓ_1 minimization and OMP, and so is attractive for solving large-scale problems.

We use phase diagrams to compare algorithm performance. The problem of recovering a k-sparse vector x_0 from (y, Φ) where Φ is random $n \times N$ and $y = \Phi x_0$ is represented by a point $(n/N, k/n)$ in this diagram; here the interesting range is $k < n < N$. For n large, StOMP correctly recovers (an approximation to) the sparsest solution of $y = \Phi x$ over a region of the sparsity/indeterminacy plane comparable to the region where ℓ_1 minimization is successful. In fact, StOMP outperforms both ℓ_1 minimization and OMP for extremely underdetermined problems. We rigorously derive a conditioned Gaussian distribution for the matched filtering coefficients at each stage of the procedure and rigorously establish a large-system limit for the performance variables of StOMP. We precisely calculate large-sample phase transitions; these provide asymptotically precise limits on the number of samples needed for approximate recovery of a sparse vector by StOMP.

We give numerical examples showing that StOMP rapidly and reliably finds sparse solutions in compressed sensing, decoding of error-correcting codes, and overcomplete representation.