CoSaMP

Iterative signal recovery
from incomplete and inaccurate samples

Joel A. Tropp

Applied & Computational Mathematics
California Institute of Technology
jtropp@acm.caltech.edu

Joint with D. Needell (UC-Davis).

Research supported in part by NSF and DARPA
The Sparsity Heuristic

A \textit{sparse signal} has fewer degrees of freedom than its nominal dimension.
Example: Wavelet Sparsity

Courtesy of J. Romberg
Example: Time–Frequency Sparsity

Data provided by L3 Communications

CoSaMP (SIAM IS08, 8 July 2008)
Quantifying Sparsity

Let \( \{\psi_k : k = 1, 2, \ldots, N\} \) be an orthobasis for \( \mathbb{R}^N \).

The coefficients of \( x \) with respect to the basis are

\[
f_k = \langle x, \psi_k \rangle \quad \text{for } k = 1, 2, \ldots, N\]

The signal is \textit{s-sparse} when \( \#\{k : f_k \neq 0\} \leq s \).

Generalization: the signal is \textit{p-compressible} with magnitude \( R \) if

\[
|f|_{(k)} \leq R \cdot k^{-1/p} \quad \text{for } k = 1, 2, \ldots, N
\]

\( p \)-compressible is slightly weaker than “in \( \ell_p \)” for each \( p > 0 \).
Approximating Compressible Signals

Consider a signal $p$-compressible w.r.t. the standard basis

$$|x|_{(k)} \leq R \cdot k^{-1/p} \quad \text{for } k = 1, 2, 3, \ldots$$

Approximating $x$ by its $s$ largest terms gives error

$$\|x - x_s\|_2 \leq R \cdot \left[ \sum_{k > s} k^{-2/p} \right]^{1/2} \approx R \cdot \left[ \int_s^\infty u^{-2/p} \, du \right]^{1/2} \approx R \cdot s^{1/2 - 1/p}$$

Compressible signals are well approximated by sparse signals

Fundamental idea behind transform coding
Consider the class of 0–1 signals in $\mathbb{R}^N$ with exactly $s$ ones.

Clearly need \textit{at least} $\log_2 \binom{N}{s}$ bits to distinguish signals.

By Stirling’s approximation, about $s \log(N/s)$ bits.

When $s \ll N$, signals contain much less information than the ambient dimension suggests.

A simple \textit{adaptive} coding scheme can achieve this rate.
A *sample* is the value of a linear functional applied to the signal

**Examples:**

- CCD: Point intensity of an image
- ADC: Voltage of an electrical signal at a point in time
- MRI: Frequency in the 2D Fourier transform of an image
- CAT: Line integral of density in one direction

Some of these technologies acquire samples in batches

We wish to acquire signals with as few samples as possible
Design linear sampling operator $\Phi : \mathbb{C}^N \rightarrow \mathbb{C}^m$

Suppose $x$ is an unknown (compressible) signal in $\mathbb{C}^N$

Collect noisy samples $u = \Phi x + e$

Problem: Given samples $u$, approximate $x$
Restricted Isometries

Abstract property of sampling operator supports efficient sampling

\[ \Phi \] has the restricted isometry property of order \( 2s \) when

\[
(1 - c) \|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + c) \|x\|_2^2 \quad \text{whenever} \quad \|x\|_0 \leq 2s
\]

\( \Phi \) preserves geometry of \( s \)-sparse signals (take \( x = y - z \))

W.h.p., a Gaussian sampling operator has RIP(\( 2s \)) when

\[
m \geq Cs \log(N/s)
\]

Gaussian matrices are practically useless

References: [Candès–Tao 2006, Rudelson–Vershynin 2006]
Practical Sampling Operators

- Partial Fourier matrices [CRT 2006]
  - Each row of $\Phi$ is chosen at random from rows of unitary DFT $\mathcal{F}_N$

- Random demodulator [Rice DSP 2006]

\[
\Phi = \begin{bmatrix}
1 & \ldots & 1 \\
1 & \ldots & 1 \\
\vdots & \ddots & \vdots \\
\end{bmatrix}_{m \times N} \begin{bmatrix}
\pm1 & \pm1 & \ldots \\
\end{bmatrix}_{N \times N}
\]

- W.h.p., both have RIP($2s$) when $m \geq Cs \log^\alpha N$

- Certain technologies can acquire these samples efficiently
- Fast matrix–vector multiplies!
Desiderata for Recovery Algorithm

- Works for general sampling schemes
- Succeeds with minimal number of samples
- Tolerates noise in samples
- Produces approximations with optimal error bound
- Yields rigorous guarantees on resource requirements
- Exploits structured sampling matrices
CoSaMP$(\Phi, u, s)$

**Input:** Sampling operator $\Phi$, noisy sample vector $u$, sparsity level $s$

**Output:** An $s$-sparse approximation $a$ of the target signal

\[ k = 0 \]
\[ a^k = 0 \] \{ Initialization \}

while halting criterion false

\[ v \leftarrow u - \Phi a^k \] \{ Update samples \}
\[ y \leftarrow \Phi^* v \] \{ Form signal proxy \}

\[ \Omega \leftarrow \text{supp}(y_{2s}) \] \{ Identification \}
\[ T \leftarrow \Omega \cup \text{supp}(a^k) \] \{ Merge supports \}

\[ b|_T \leftarrow \Phi_T^* u \] \{ Signal estimation by least squares \}
\[ b|_{T^c} \leftarrow 0 \]

\[ a^{k+1} \leftarrow b_s \] \{ Prune to obtain next approximation \}
\[ k \leftarrow k + 1 \]
end while

\[ a \leftarrow a^k \] \{ Return final approximation \}

CoSaMP (SIAM IS08, 8 July 2008) 13
Cost per Iteration

- Update samples and form signal proxy:
  \[ y \leftarrow u - \Phi a^k \quad \text{and} \quad y \leftarrow \Phi^* v \]

- One matrix–vector multiplication each

- Signal approximation by least squares:
  \[ b_T \leftarrow \Phi_{T}^\dagger u \]

- Use conjugate gradient to apply pseudoinverse
- Each iteration requires two matrix–vector multiplies
- Assuming RIP(2s), constant number of iterations for fixed accuracy

- Constant number of matrix–vector multiplies per CoSaMP iteration!
Theorem 1. [CoSaMP] Suppose that

- the sampling matrix $\Phi$ has RIP($2s$),
- the sample vector $u = \Phi x + e$,
- $\eta$ is a precision parameter,
- $\mathcal{L}$ bounds cost of a matrix–vector multiply with $\Phi$ or $\Phi^*$.

Then CoSaMP produces a $2s$-sparse approximation $a$ such that

$$\|x - a\|_2 \leq C \max \left\{ \eta, \frac{1}{\sqrt{s}} \|x - x_s\|_1 + \|e\|_2 \right\}$$

with execution time $O(\mathcal{L} \cdot \log(\|x\|_2 / \eta))$.

Need $m \geq Cs \log^\alpha N$ samples for restricted isometry hypothesis
Corollary 2. [Compressible signals] Suppose

- the sampling matrix $\Phi$ has RIP($2s$),
- the signal $x$ is $p$-compressible with magnitude $R$,
- the sample vector $u = \Phi x + e$,
- $\mathcal{L}$ bounds cost of a matrix–vector multiply with $\Phi$ or $\Phi^*$.

Then CoSaMP produces a $2s$-sparse approximation $a$ such that

$$\| x - a \|_2 \leq C \left[ Rp^{-1} \cdot s^{1/2-1/p} + \| e \|_2 \right]$$

with execution time $O(\mathcal{L} \cdot p^{-1} \log s)$. 

CoSaMP (SIAM IS08, 8 July 2008)
To learn more...

E-mail:

- jtropp@acm.caltech.edu
- dneedell@math.ucdavis.edu

Web: http://www.acm.caltech.edu/~jtropp

Relevant Papers:

- NTV, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” accepted to ACHA
- T and Rice DSP, “Beyond Nyquist: Efficient sampling of sparse, bandlimited signals,” in preparation
- N and Vershynin, “Stable signal recovery from incomplete and inaccurate samples,” submitted
- T and Gilbert, “Signal recovery from random measurements via Orthogonal Matching Pursuit,” Trans. IT, Dec. 2007.