ECE 8201: Low-dimensional Signal Models for High-dimensional Data Analysis

Yuejie Chi
Departments of ECE and BMI
The Ohio State University

September 24, 2015

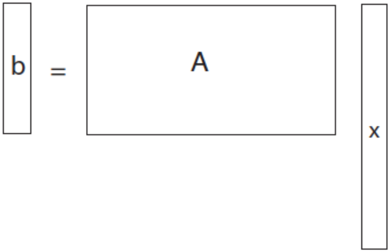
THE OHIO STATE UNIVERSITY

Time, location, and office hours

- **Time:** Tue/Thu 5:30-6:55pm
  - No lectures on 10/1 (Thu) and 11/10 (Tue) due to travel of the instructor. We will extend each lecture by 5 minutes to compensate these two lectures.
- **Location:** Bolz Hall 314
- **Instructor:** Dr. Yuejie Chi (chi.97@osu.edu)
- **Office:** 606 Dreese Lab
- **Office hours:** Thu: 3:00-5:00pm.

Underdetermined linear systems

We're interested in solving *underdetermined systems of linear equations*.



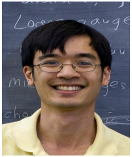
- Estimate $\mathbf{x} \in \mathbb{R}^n$ from linear measurements $\mathbf{b} = \mathbf{A}\mathbf{x} \in \mathbb{R}^m$, where $m \ll n$.
- Seems to be hopelessly ill-posed, since more unknowns than equations...

Compressed Sensing

Compressed Sensing/Compressive Sensing/Compressive Sampling allows perfect recovery of underdetermined linear systems by exploiting additional structures in **x**, e.g. sparsity, low-rankness, etc.

- Compressed Sensing [Name coined by David Donoho] was pioneered by Donoho and Candès, Tao and Romberg in 2004.






Compressed sensing
DL Donoho - Information Theory, IEEE Transactions on, 2006 - ieeexplore.ieee.org
Abstract—Suppose is an unknown vector in(a digital image or signal); we plan to measure general linear functionals of and then reconstruct. If is known to be compressible by transform coding with a known transform, and we reconstruct via the nonlinear procedure ...
Cited by 9878 Related articles All 31 versions Cite Save More

Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information
EJ Candès, J Romberg, T Tao - Information Theory, IEEE ..., 2006 - ieeexplore.ieee.org
Abstract—This paper considers the model problem of recon-structing an object from incomplete frequency samples. Consider a discrete-time signal and a randomly chosen set of frequencies. Is it possible to reconstruct from the partial knowledge of its Fourier ...
Cited by 6692 Related articles All 38 versions Cite Save

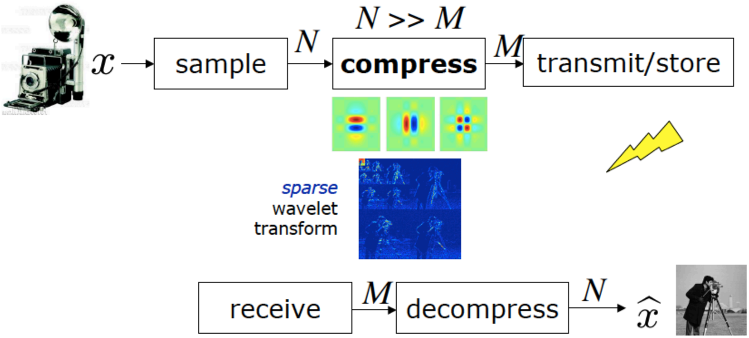
Three motivating examples

Through three motivating examples, we will see such underdetermined linear systems arise quite frequently in science and engineering applications, and exhibit intriguing signal structures that make the problem well-posed and solvable using, e.g., convex/non-convex optimization techniques.

- Sparse signal recovery
- Low-rank matrix completion
- Phase retrieval

Example 1: Sparse signal recovery

Conventional paradigm of data acquisition: acquire then compress.



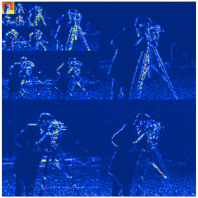
Why can we compress? There's no loss in quality between the raw image and its JPEG compressed form.

Sparse representation

- **Sparsity:** Many real world signals admit *sparse representation.* The signal $\mathbf{s} \in \mathbb{C}^n$ is sparse in a basis $\mathbf{\Psi} \in \mathbb{C}^{n \times n}$, as

$$\mathbf{s} = \mathbf{\Psi}\mathbf{x};$$

- Images are sparse in the wavelet domain.



- The number of large coefficients in the wavelet domain is small, which allows *compression*.

Sparse representation allows compression

Take a mega-pixel image

- Compute 10^6 wavelet coeffients
- Set to zero all but the 2.5×10^4 largest coefficients
- Invert the wavelet transform

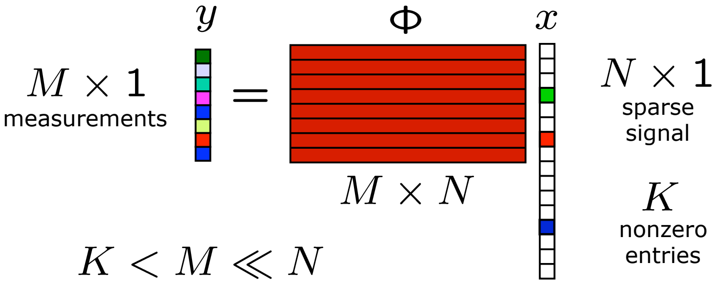
1 megapixel image

25k term approximation

Compressed sensing: compression on the fly

- Why cannot we directly the *compressed* data and then reconstruct?

$$y_i = \langle \mathbf{a}_i, \mathbf{x} \rangle, \quad i = 1, \ldots, m$$

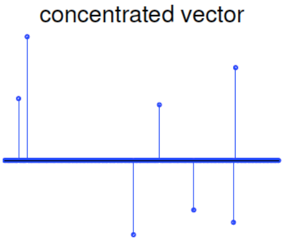


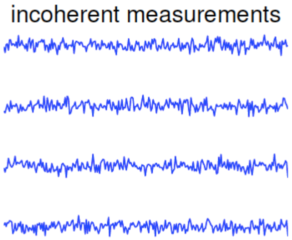
- Mathematically, this give rises to an underdetermined system of equations, where the signal of interest is *sparse*.

Sparse recovery

Questions of interest include:

- How to design the measurement/sampling matrix?
- What are the efficient algorithms?
- Are they stable with respect to noise?
- How many measurements are necessary/sufficient?

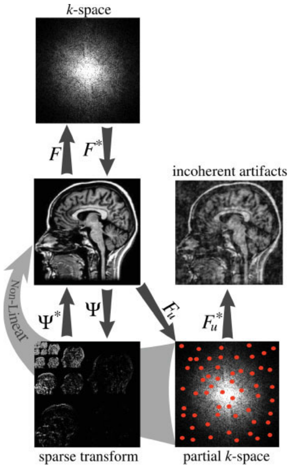




Spoiler: It turns out $m \gtrsim K \log n$ random measurements will suffice.

CS application in MRI

- Imaging speed is important in many MRI applications, and it is of great importance to reduce the image acquisition time;
- MR images are acquired by sampling the *k*-space (frequency domain) - take a long time to fully sample it!
- By subsampling the *k*-space, the acquisition time is reduced; however image artifacts arise if using conventional reconstruction.
- Novel algorithms that exploit sparsity allows much better reconstruction.



Transform Domain Sparsity of MRI

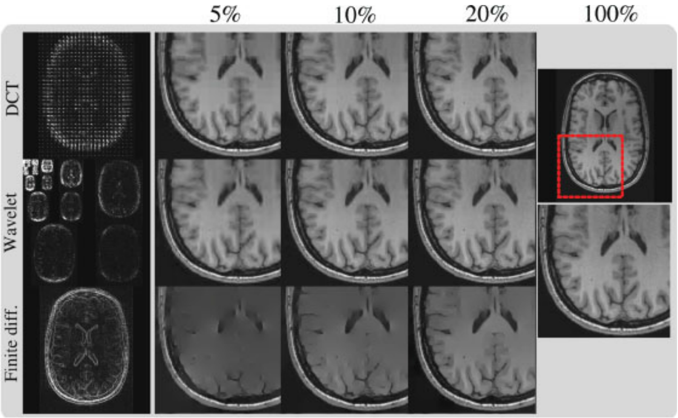


Figure : Transform-domain sparsity of images. The DCT, wavelet, and finite-differences transforms were calculated for the image (Left column). The image was then reconstructed from a subset of 5, 10, and 20% of the largest transform coefficients [Lustig et. al. 2007].

Accelerate MRI via CS Reconstruction

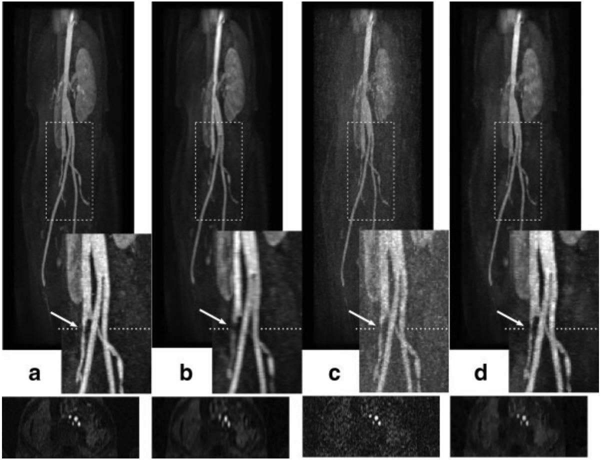


Figure : Reconstruction from 5-fold accelerated acquisition of first-pass contrast enhanced abdominal angiography. (a) Reconstruction from complete data; (b) LR; (c) ZF-w/dc; (d) CS reconstruction from random subsampling [Lustig et. al. 2007].

Many applications of CS

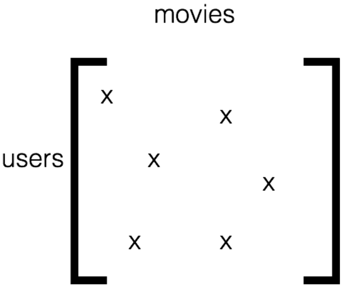
CS is expected to become useful when measurements are

- expensive (e.g. fuel cell imaging, near IR imaging)
- slow (e.g. MRI)
- beyond current capabilities (e.g. wideband analog to digital conversion)
- wasteful
- missing
- etc..

Example 2: The Netflix problem

In 2006, Netflix offered a $1 million prize to improve its movie rating prediction algorithm.

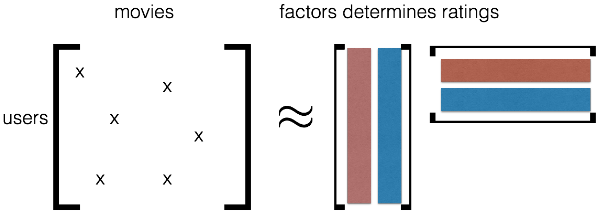
- How to estimate the missing ratings?



- About a million users, and 25,000 movies, with sparsely sampled ratings

Low-rank matrix completion

- Completion problem: consider $\mathbf{M} \in \mathbb{R}^{n_1 \times n_2}$ to represent Netflix data, we may model it through factorization:



- In other words, the rank r of $\mathbf{M}$ is much smaller than its dimension $r \ll \min\{n_1, n_2\}$.

Low-rank matrix completion

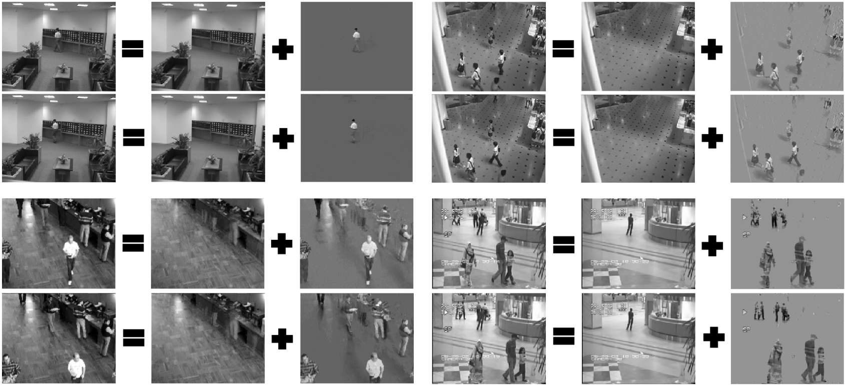
Questions of interest include:

- What are the efficient algorithms for matrix completion?
- Are they stable with respect to noise and outliers?
- How many measurements are necessary/sufficient?

Spoiler: Under appropriate conditions, matrix completion is possible from $m \gtrsim r \max\{n_1, n_2\} \log^2(\max\{n_1, n_2\})$ samples.

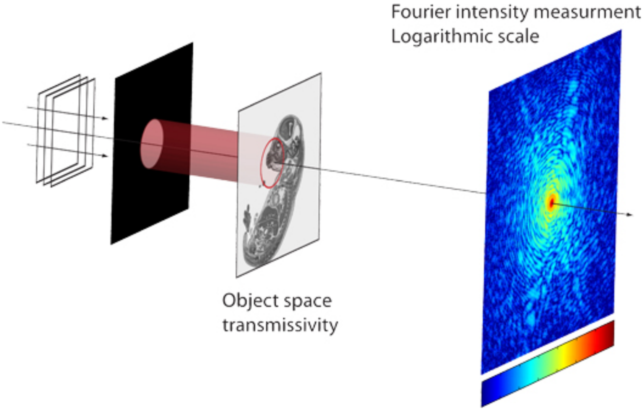
Applications in computer vision

Separation of background (low-rank) and foreground (sparse) in video:



Example 3: Phase retrieval

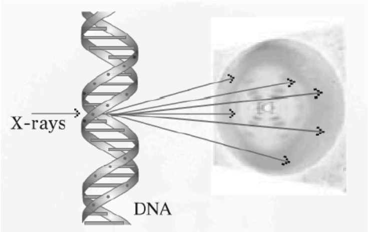
In optics, we only observe the magnitude of measurements but not the phase.



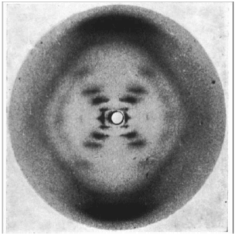
What can we tell about the structure from the intensity of its Fourier transform?

X-rays crystallography

- Useful in many applications, such as X-rays crystallography, which allows determination of atomic structures within a crystal
- Example: discovery of Double-Helix structure of DNA (1962 Nobel Prize)



principle



Franklin's photograph

- 10 Nobel Prizes in X-ray crystallography so far

Phase retrieval

We formulate the phase retrieval problem as follows.

- Let the measurements vectors be $\{\mathbf{a}_i \in \mathbb{R}^n\}_{i=1}^m$.
- For each i, we measure

$$y_i = |\langle \mathbf{a}_i, \mathbf{x} \rangle|^2, \quad i = 1, \ldots, m.$$

- How to recover $\mathbf{x} \in \mathbb{R}^n$?
- Standard Gerchberg-Saxton (or Fienup) iterative algorithms suffer from local minima.

Note these equations are *quadratic* in $\mathbf{x}$, and no sparsity is assumed for $\mathbf{x}$, so $m \geq n$ would be necessary.

Applying lifting

- Expand

$$\begin{aligned} y_i &= |\langle \mathbf{a}_i, \mathbf{x} \rangle|^2 \\ &= (\mathbf{a}_i^* \mathbf{x})(\mathbf{a}_i^* \mathbf{x})^* \\ &= \mathbf{a}_i^* (\mathbf{x}\mathbf{x}^*) \mathbf{a}_i := \mathbf{a}_i^* \mathbf{X} \mathbf{a}_i, \end{aligned}$$

  where $\mathbf{X} = \mathbf{x}\mathbf{x}^*$ is a rank-one symmetric matrix.
- To recover $\mathbf{x}$, it is equivalent to recovering $\mathbf{X}$ (up to rotational ambiguity).
- We again convert the problem to an underdetermined linear system!

Phase retrieval

Questions of interest include:

- How to design the measurement/sampling matrix?
- What are the efficient algorithms?
- Are they stable with respect to noise?
- How many measurements are necessary/sufficient?

Spoiler: Under appropriate conditions, phase retrieval is possible from $m \gtrsim n$ samples.

Variations on a common theme

- We can solve an underdetermined linear system if one has more information about the underlying signal, together with a well-posed sampling scheme.
- We will show for the three examples above, there is a convex optimization algorithm that provides *near-optimal* performance guarantees.
- We will also discuss extensions of the theory:
  - What if my measurements are nonlinear?
  - What if I have some low-dimensional signal structure that is not sparse/low-rank?

If time allows...

- Dimensionality reduction
- Dictionary learning
- Online learning

Warning

- There will not be textbooks, only references...
  - Boyd and Vandenberghe, *Convex Optimization*, Available online.
  - Foucart and Rauhut, *A Mathematical Introduction to Compressive Sampling*. Available online.
  - Vershynin, *Introduction to the non-asymptotic analysis of random matrices.* Available on arxiv:1011.3027.
  - We will also provide references on related papers in the literature.
- There will be quite a few PROOFS....
- This course is research-oriented.

Prerequisites

- Probabilty
- Linear algebra
- Knowledge in convex optimization is a plus
- Mathematical maturity
- Know how to use MATLAB

Grading

- Homework: 25%
  - We shall have no more than five homeworks.
- Midterm paper: 25% (Due at the beginning of Week 9)
  - Literature review that sets the stage for the final project.
- Final Project: 50% (in-class presentation 15% + final report 35%)
  - Around week 12, we will ask for a short project proposal/progress report to make sure you're on the right track.
  - In-class presentation during Week 16, with the final report due at the end of Week 16.

Projects

- Projects should be independent work.
- Topics are flexible: it can be related to your current research, but use something learned from this course.
- A good project should contain something novel and useful:
  - it could be applications of existing algorithms to a new domain;
  - it could be proposal of a new algorithm;
  - it could be new theoretical analysis of an existing algorithm;
  - etc...

  It is important that you justify its novelty!

Project evaluation

- In-class presentation
  - Each student will be allocated 15 minutes - this is about the time for a conference presentation.
- Final report
  - The final report should follow the format of an IEEE journal paper. (It could become a publication of yours!)

How to best prepare for the lectures

Read, read, read!

- Chapter 1-2 [Foucart-Rauhut]
- *Introduction to Compressed Sensing*
  Book chapter by Davenport et.al.
  `http://statweb.stanford.edu/~markad/publications/ddek-chapter1-2011.pdf`