Compressed sensing and imaging

Anwei Chai

Computational and Mathematical Engineering
Stanford University

MATH00112230: Selected Topics in Advanced Statistics

Joint with Laurent Demanet (MIT), George Papanicolaou (Stanford)

Outline

- Part I. Solving linear equations
  - least squares
  - regularization (ℓ_2 penalty method)
  - regularization (smoothness)
  - regularization (sparsity ℓ_1)
- Part II. Imaging localized scatterers
  - Array imaging
  - Kirchhoff migration, least square imaging
  - Multiple signal classification (MUSIC)
  - Singular value decomposition and ℓ_1 based method

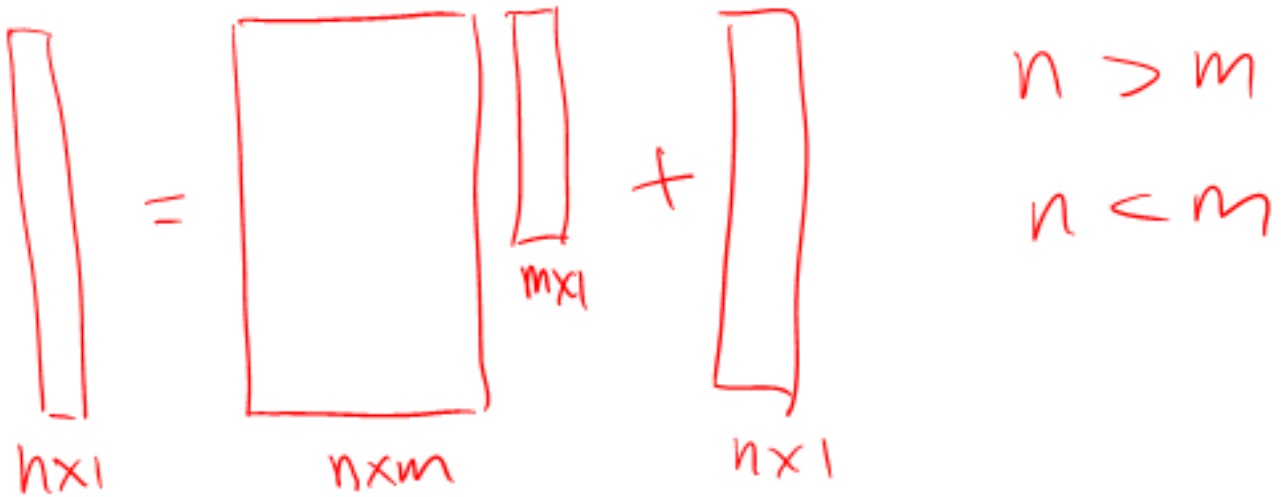
Part I. Solving linear equations – from least square to compressed sensing

Solving linear equations

- Model:

$$\mathbf{y} = A\mathbf{x} + \mathbf{n} \tag{1}$$

- $\mathbf{y}$ is data, $\mathbf{x}$ is model image, $\mathbf{n}$ is noise



Solving linear equations – least square

- $n > m$ (over-determined): Find minimizer of $J(\mathbf{x}) = \|A\mathbf{x} - \mathbf{y}\|_2^2$
- Solve $\nabla J(\mathbf{x}) = A^*(A\mathbf{x} - \mathbf{y}) = 0$ for minimizer
  - If A^*A is invertible, $\widehat{\mathbf{x}} = A^\dagger \mathbf{y}$ and $A^\dagger = (A^*A)^{-1}A^*$ (pseudo-inverse)
  - If A^*A is not invertible ($\ker A \neq 0$), $\widehat{\mathbf{x}} = A^\dagger \mathbf{y} + \mathbf{z}$, $\forall \mathbf{z} \in \ker A$
  - Calculate $A^\dagger$ from singular value decomposition (SVD)
- $n < m$ (under-determined): infinitely many solutions if any
- least square gives least norm solution

$$\widehat{\mathbf{x}} = A^\dagger \mathbf{y}, \quad A^\dagger = A^*(AA^*)^{-1}$$

$J(x) = \|x\|_2^2 + \lambda^*(Ax - y)$

Solving linear equations – ℓ_2 penalty

- Motivation: consider $n = m$ and (1) has solution

$$\widehat{\mathbf{x}} = A^{-1}\mathbf{y} = \mathbf{x} + A^{-1}\mathbf{n}$$

$y = Ax + n$

$\frac{1}{\epsilon}$

bad when A has small singular values

- Assume $\|\mathbf{x}\|_2$ small

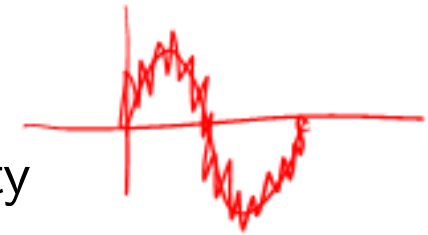
$$\min \|\mathbf{x}\|_2 \quad \text{s.t. } \|A\mathbf{x} - \mathbf{y}\|_2 \leq \varepsilon \tag{2}$$

- (2) has several equivalent forms

$$\min \|A\mathbf{x} - \mathbf{y}\|_2 \quad \text{s.t. } \|\mathbf{x}\|_2 \leq \varepsilon'$$
$$\min \|\mathbf{x}\|_2^2 + \lambda \|A\mathbf{x} - \mathbf{y}\|_2^2$$
$$\min \|A\mathbf{x} - \mathbf{y}\|_2^2 + \lambda' \|\mathbf{x}\|_2^2$$

- Solution to Lagrangian form of (2) is $\widehat{\mathbf{x}} = (A^*A + \lambda I)^{-1} A^*\mathbf{y}$

Solving linear equations – smooth solution



- Assume $\|B\mathbf{x}\|_2$ small: variation of ℓ_2 penalty

$$\text{diag}(B,0) = 1,\ \text{diag}(B,1) = -1,\ \text{diag}(B,k) = 0,\ \ k \neq 0,1$$

- $B\mathbf{x}$ is finite difference approximation to derivative
- Tikhonov regularization in image processing and Hodrick-Prescott filtering in time series analysis
- Solution is given by $\widehat{\mathbf{x}} = (A^*A + \lambda B^*B)^{-1}A^*\mathbf{y}$
- View point from numerical linear algebra: add extra term (λI or λB^*B) to make A^*A well conditioned to take inverse

Solving linear equations – ℓ_1 penalty

- Assume under-determined A and sparse structure of $\mathbf{x}$, i.e. $\|\mathbf{x}\|_0 = |\operatorname{supp} \mathbf{x}| = |\{i : x_i \neq 0\}|$ small
- When $A = I$, choose $\mathbf{x}$ s.t.

$$\min \|\mathbf{x}\|_0 \quad \text{s.t. } \|A\mathbf{x} - \mathbf{y}\|_2 \leq \sigma \tag{3}$$

- When $A \neq I$, very difficult to identify nonzeros in $\mathbf{x}$
- Relaxation to ℓ_1 norm of $\mathbf{x}$: Basis Pursuit Denoising (BPD) or Least Angle Shrinkage and Selector Operator (LASSO)

$$\min \|\mathbf{x}\|_1 \quad \text{s.t. } \|A\mathbf{x} - \mathbf{y}\|_2 \leq \sigma \tag{4}$$

- Solution is given by soft-threshold

$$\widehat{x}_i = \mathcal{S}_\tau y_i = \begin{cases} y_i - \tau \operatorname{sgn} y_i, & |y_i| \geq \tau \\ 0, & \text{otherwise.} \end{cases}$$

Solving linear equations – connection with uncertainty principle

- X is transformed $\mathbf{x}$ (e.g. in frequency domain), $N_t = |\operatorname{supp} \mathbf{x}|$, $N_\omega = |\operatorname{supp} X|$
  - $N_t N_\omega \geq m$, $N_t + N_\omega \geq 2\sqrt{m}$ (Donoho & Stark 1989)
  - $N_t + N_\omega \geq m + 1$ when m is prime (Tao 2003)
- Let $A = [I, F]$ where $m = 2n$ and F is normalized Fourier matrix; $T = \operatorname{supp} \mathbf{x} \cap \{1, 2, \ldots, n\}$, $W = \operatorname{supp} \mathbf{x} \cap \{n+1, n+2, \ldots, 2n\}$. Assume there is no noise (Donoho & Huo 1999)
  - $|T| + |W| < \sqrt{n}$, solution of (3) is unique and equals $\mathbf{x}$
  - $|T| + |W| < \sqrt{n}/2$, solution of (4) is unique and equals $\mathbf{x}$
  - $\exists n, \sqrt{n} \in \mathbb{Z}$ and $|T| + |W| = \sqrt{n}$, solutions of (3) and (4) are not unique

if $|T|$ small, no small number of frequencies can recover $\mathbf{x}$

Solving linear equations – compressed sensing

- $A = I$: soft-threshold
- $A = [I, F]$: uncertain principle
- $A = [\Phi_1, \Phi_2]$: mutual incoherence
  $\mu = \max\{|\langle \phi_1, \phi_2 \rangle|, \phi_1 \in \Phi_1, \phi_2 \in \Phi_2\}$ and $\|\mathbf{x}\|_0 < (1 + 1/\mu)/2$ (also true for sparse solution under general A)
- Compressed sensing: sample enough data
  - Counting faces of random polytope (Donoho 2004, Donoho & Tanner 2009)
  - Restricted Isometry Property (RIP) (Candés & Tao 2004-2005)

Matrix A is called to satisfy RIP-S if for any subset $T \in \{1, 2, \ldots, m\}$, $|T| = S$, there exists $0 < \delta_S < 1$ such that

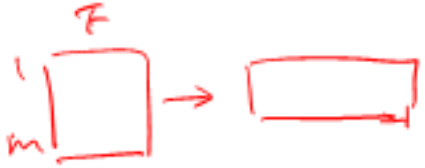
$$(1 - \delta_S)\|\mathbf{c}\|_2^2 \leq \|A_T \mathbf{c}\|_2^2 \leq (1 + \delta_S)\|\mathbf{c}\|_2^2$$

for any $\mathbf{c} \in \mathbb{R}^S$

Solving linear equations – compressed sensing

Three kinds of matrices are proven to satisfy exact reconstruction using (4) with high probability

- uniformly randomly sampled Fourier matrix



Choose $n \geq 22(1+\delta)\|\mathbf{x}\|_0 \log m$ rows uniformly at random from Fourier matrix F. Then (4) recovers $\mathbf{x}$ with probability $p = 1 - \mathcal{O}(m^{-\delta})$

- Gaussian ensemble: A_{ij}'s are iid Gaussian random variables up to normalization constant

If entries of A are sampled from $\mathcal{N}(0, 1/n)$ and $S \leq Cn/\log(m/n)$, then reconstruction succeeds with probability $p = 1 - \mathcal{O}(e^{-\gamma m})$ for some $\gamma > 0$

- Bernoulli ensemble: A_{ij}'s are iid Bernoulli random variables up to normalization constant

Same as Gaussian ensemble with A sampled from symmetric Bernoulli $\pm 1/\sqrt{n}$

Solving linear equations – compressed sensing

Deterministic exact and stability results are proven under RIP

Assume there is no noise and $\mathbf{x}$ is S sparse and $\delta_{2S} + \delta_{3S} < 1$. Then solving (4) gives exact solution.

Consider problem (1). If $\delta_{3S} + 3\delta_{4S} < 2$, then solution $\mathbf{x}^\star$ to (4) recovers original solution $\mathbf{x}$ satisfying $\|\mathbf{x}\|_0 \leq S$ up to

$$\|\mathbf{x}^\star - \mathbf{x}\|_2 \leq C_S \sigma$$

If $\mathbf{x}_S$ is the approximation to any $\mathbf{x}$ by keeping only the first S largest values. Then solution $\mathbf{x}^\star$ to (4) gives approximation

$$\|\mathbf{x}^\star - \mathbf{x}\|_2 \leq C_S \sigma + C_S' \frac{\|\mathbf{x} - \mathbf{x}_S\|_1}{\sqrt{S}}$$

Solving linear equations – numerical solution

- Off-shelf: any optimization package
  - linear programming for equality constraint
  - second order cone programming for inequality constraint
  - semi-definite programming for higher order problem

  "easy" to implement but extremely inefficient for large scale problem (good for problem of size less than 100)
- iterative methods:
  - Iterative threshold method (IST)
  - Bregman, linearized Bregman
  - Message passing method

  straightforward to implement, fast, and efficient for large scale problem

Solving linear equations – application

- error correction in communication
- mathematical image processing (motion correction)
- medical imaging (MRI to reduce treatment time expense)
- statistics (feature selection)
- a lot more, see list at *http://dsp.rice.edu/cs/*

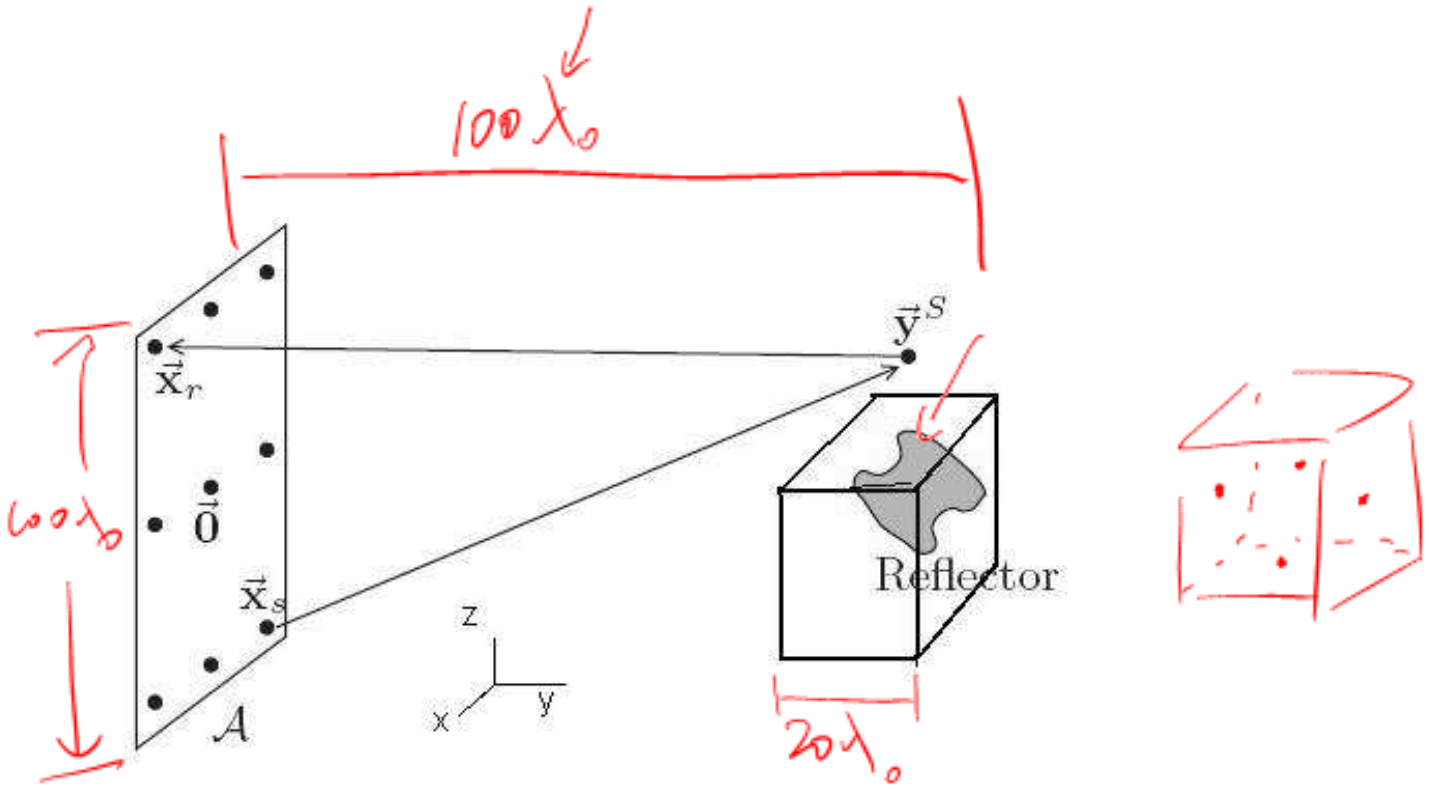
Part II. Imaging localized scatterers

Overview – questions and methods

- Compare several imaging methods regarding resolution and stability, i.e. loss of resolution when signal to noise ratio (**SNR**) is low at imaging array or medium between array and targets is inhomogeneous
- Methods to be compared are

  KM: travel time migration imaging
  LSQ: least square imaging
  MUSIC: MUlitiple SIgnal Classification
  CS: ℓ_1 minimization with or without SVD

Array imaging – schematic



Configuration: range is $100\lambda_0$; array size is between $100\lambda_0$ and $200\lambda_0$ with transducers half of wavelength apart; target area is box of length $20\lambda_0$ with half of wavelength resolution

Array imaging – formulation

Time harmonic Green function in homogeneous medium with speed c_0 is

$$\widehat{G}_0(\mathbf{x}, \mathbf{y}, \omega) = \frac{e^{i\frac{\omega}{c_0}|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}$$

Response at each sensor $\mathbf{x}_r$ on the array excited by the signal emitted from $\mathbf{x}_s$ and reflected by scatters of reflectivity ρ_j, $j = 1, \ldots, M$ is given by

$$\widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega) = \sum_{j=1}^{M} \rho_j \widehat{G}_0(\mathbf{x}_r, \mathbf{y}_j, \omega)\widehat{G}_0(\mathbf{x}_s, \mathbf{y}_j, \omega).$$

Let $\widehat{g}_0(\mathbf{y}, \omega) = [\widehat{G}_0(\mathbf{x}_r, \mathbf{y}, \omega)]$ and with excitation $\mathbf{f}(\omega)$,

$$[(\widehat{g}_0^T(\mathbf{y}, \omega)f(\omega))\widehat{g}_0(\mathbf{y}, \omega)]\boldsymbol{\rho} = \widehat{\Pi}(\omega)\mathbf{f}(\omega)$$

Array imaging – formulation

Rewriting by letting $\mathcal{A}_f(\omega) = [(\widehat{g}_0^T(\mathbf{y},\omega)f(\omega))\widehat{g}_0(\mathbf{y},\omega)]$, reflectivity vector ρ satisfies

$$\mathcal{A}_f(\omega)\boldsymbol{\rho} = \widehat{\Pi}(\omega)\mathbf{f}(\omega) \tag{5}$$

Imaging is inverse problem which needs to solve (5) (with or without subsampling)

Kirchhoff migration and least square imaging

- **LSQ**: $\min \|\boldsymbol{\rho}\|_2$ s.t. $\mathcal{A}_f(\omega)\boldsymbol{\rho} = \widehat{\Pi}(\omega)\mathbf{f}(\omega)$ and solution is given by $\widehat{\boldsymbol{\rho}}_{LSQ} = \mathcal{A}^{\dagger}(\omega)\widehat{\Pi}(\omega)\mathbf{f}(\omega)$
- **KM**: approximate solution $\boldsymbol{\rho}_{KM} = \mathcal{A}_f^H(\omega)\widehat{\Pi}(\omega)\mathbf{f}(\omega)$
- **KM** only gives the location of scatterers; **LSQ** computes reflectivity of scatterers due to additional terms in $\mathcal{A}^{\dagger}$ besides $\mathcal{A}^*$.

Multiple Signal Classification (MUSIC)

MUSIC assumes data are formed by some known basis. In imaging, it means data are formed by Green function vectors $\widehat{g}_0(\mathbf{y}, \omega)$. Let SVD of response matrix $\widehat{\Pi}(\omega)$ be $\widehat{\Pi}(\omega) = U\Sigma V^* = \sum_{j=1}^{L} \sigma_j \mathbf{u}_j \mathbf{v}_j^H$

- Can show $L \approx M$, $\mathbf{u}_j \doteq \frac{\widehat{g}_0(\mathbf{y}_{k_j}, \omega)}{\|\widehat{g}_0(\mathbf{y}_{k_j}, \omega)\|_2}$, $\mathbf{v}_j \doteq \overline{\mathbf{u}_j}$, $\sigma_j \doteq \rho_j \|\widehat{g}_0(\mathbf{y}_{k_j}, \omega)\|_2^2$
- Assume $\widehat{g}_0(\mathbf{y}, \omega)$'s are normalized, $\mathbf{u}_k^* \widehat{g}_0(\mathbf{y}^S, \omega)$ has peak when $\mathbf{y}^S$ is close to a scatterer
- **MUSIC** imaging is given by

$$\widehat{\rho}_{MUSIC}(\mathbf{y}^S) = \frac{\min_{\mathbf{y}^S} \mathcal{G}_{MUSIC}(\mathbf{y}^S)}{\mathcal{G}_{MUSIC}(\mathbf{y}^S)},$$

where

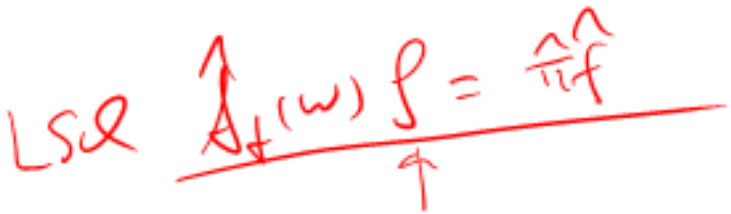
$$\mathcal{G}_{MUSIC}(\mathbf{y}^S) = \|\mathcal{P}_N \widehat{g}_0(\mathbf{y}^S, \omega)\|_2^2$$

and

$$\mathcal{P}_N \widehat{g}_0(\mathbf{y}^S, \omega) = \sum_{j=1}^{M} \left(\mathbf{u}_j^H \widehat{g}_0(\mathbf{y}^S, \omega) \right) \mathbf{u}_j - \widehat{g}_0(\mathbf{y}^S, \omega)$$

Multiple Signal Classification (MUSIC)

MUSIC only gives location of scatterers in terms of the peak values of the functional. Solving resulting overdetermined least square problem by columns of $\mathcal{A}_f(\omega)$ corresponding to large **MUSIC** functional



Imaging with ℓ_1 optimization

- Solution $\boldsymbol{\rho}$ is sparse
- Application of compressed sensing directly

$$\min \|\boldsymbol{\rho}\|_1 \quad \text{s.t. } \Phi \mathcal{A}_f(\omega)\boldsymbol{\rho} = \Phi \widehat{\Pi}(\omega)\mathbf{f}(\omega)$$

  Φ is uniform random sampling operator
- Accurate results subject to high probability of success and hard to determine the sampling ratio
- Physical experiment by group in Georgia Tech (Gurbuz, McClellan, Scott 2009); theoretical analysis by (Fannjiang 2009)

Imaging with SVD and ℓ_1 optimization

- Implementation of random sampling, amount of subsampling, flexibility of illumination (i.e. $\mathbf{f}(\omega)$)
- SVD of response matrix is powerful to identify number of scatterers and robust to noise
- Green function vectors $\widehat{g}_0(\mathbf{y}, \omega)$ are asymptotically orthogonal, i.e.

$$\widehat{g}_0^H(\mathbf{y}_i, \omega)\widehat{g}_0(\mathbf{y}_j, \omega) = \begin{cases} \|\widehat{g}_0(\mathbf{y}_i, \omega)\|_2^2, & \text{if } i = j, \\ 0 \text{ (approximately)}, & \text{otherwise,} \end{cases}$$

provided $\lambda \ll \|\mathbf{y} - \mathbf{y}^S\|_2 \ll L$

Imaging with SVD and ℓ_1 optimization

Use SVD to reduce dimension (subsample) of data

- Project to space spanned by vectors according to nonzero singular values

$$\mathbf{u}_k^*(\omega)\mathcal{A}_f(\omega)\rho = \mathbf{u}_k^*(\omega)\widehat{\Pi}(\omega)\mathbf{f}(\omega), \quad k = 1, \ldots, M$$

- Apply illumination vectors $\mathbf{v}_k(\omega)$

$$\mathbf{u}_k^*(\omega)\mathcal{A}_{\mathbf{v}_k(\omega)}(\omega)\rho = \mathbf{u}_k^*(\omega)\widehat{\Pi}(\omega)\mathbf{v}_k(\omega) = \sigma_k, \quad k = 1, \ldots, M$$

- Rewrite new linear system as $\mathcal{B}\boldsymbol{\rho} = \mathbf{b}$

$$\mathcal{B}_{ij} = (\widehat{g}_0^T(\mathbf{y}_j, \omega)\mathbf{v}_i(\omega))\overline{\widehat{g}_0^*(\mathbf{y}_j, \omega)\mathbf{u}_i(\omega)} = \left(\mathbf{u}_i^*\widehat{g}_0(\mathbf{y}_j, \omega)\right)^2$$

$$\mathbf{b} = [\sigma_1, \cdots, \sigma_M]^T$$

Imaging with SVD and ℓ_1 optimization

- Lower-bound: reduce dimension of data to minimum
- Optimality: use singular vector as illumination
- Guaranteed solution in specific regime ($\lambda \ll \|\mathbf{y} - \mathbf{y}^S\|_2 \ll L$):
  - $\mathcal{B} = [I + E \; S]$ is full row-rank and contains an identity matrix up to normalization
  - If $\|E\|_1 \leq \varepsilon$ and $\|S\|_1 \leq 1 - \varepsilon$, then solving

$$\min \|\boldsymbol{\rho}\|_1 \quad \text{s.t. } \mathcal{B}\boldsymbol{\rho} = \mathbf{b}$$

    gives the exact solution
  - Proof relies on the fact that the solution to (4) is unique if there exists vector $\mathbf{w}$ such that inner product of $\mathbf{w}$ and any column vector in $\mathcal{B}$ is no greater than 1 (Candés & Tao 2006)
  - No specific numerical method is need since problem size becomes small and $\mathcal{B}$ is approximately sparse

Imaging with additive noise

- Use inequality constraint $\|\mathcal{A}_f(\omega)\boldsymbol{\rho} - \widehat{\Pi}(\omega)\mathbf{f}(\omega)\|_2 \leq \varepsilon$ in **LSQ** and **CS**
- Noise is generated on imaging array in homogeneous medium and only affect high frequency data, i.e. space corresponding to zero/small singular values (Borcea, Papanicolaou, Vasquez 2008)

$$\widehat{\Pi}(\omega) + \Delta\widehat{\Pi}(\omega) = [U_1 \quad U_2 U_{\Delta\widehat{\Pi}}] \begin{bmatrix} \Sigma & \\ & \Sigma_{\Delta\widehat{\Pi}} \end{bmatrix} \begin{bmatrix} V_1^* \\ V_{\Delta\widehat{\Pi}}^* V_2^* \end{bmatrix}$$

- SVD filters out noisy terms in ℓ_1 optimization; no more additional denoising process is needed

Numerical simulation

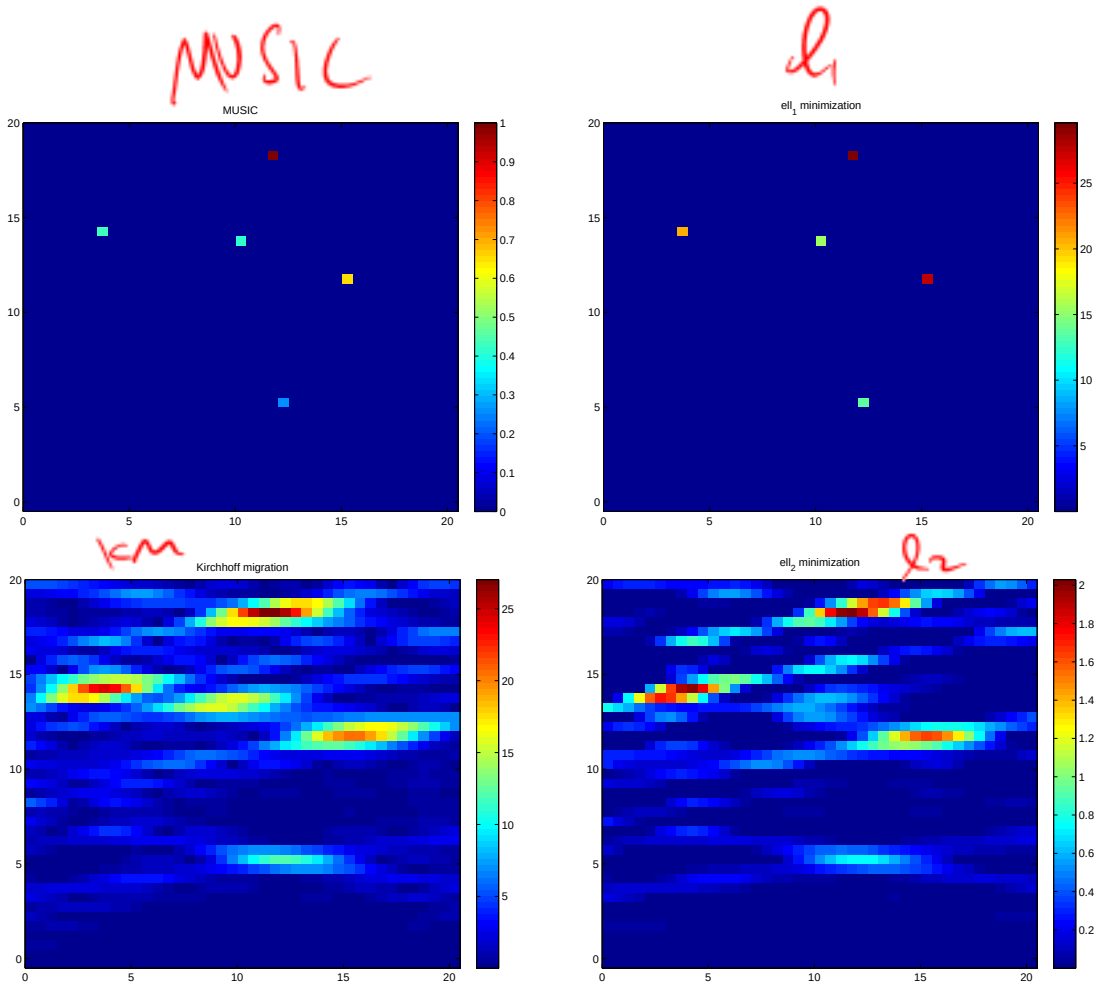
- Setup of simulation

	two dimension	three dimension
length of array	$100\lambda_0$	$100\lambda_0$
distance between detectors h	$\lambda_0/2$	$4\lambda_0$
array to IW distance L	$100\lambda_0$	$100\lambda_0$
size of IW	$20\lambda_0 \times 20\lambda_0$	$20\lambda_0 \times 5\lambda_0 \times 20\lambda_0$
sampling rate within IW k	$\lambda_0/2$	$\lambda_0/2$

The size of $\mathcal{A}_f$ is 201×1681 in two dimension and 676×18491 in three dimension

- Perturbation to $\widehat{\Pi}(\omega)$ is simulated by generating iid complex $\mathcal{N}(0, \delta p_{avg})$ and signal-to-noise ratio (SNR) is $-10 \log_{10} \delta$

Numerical simulation

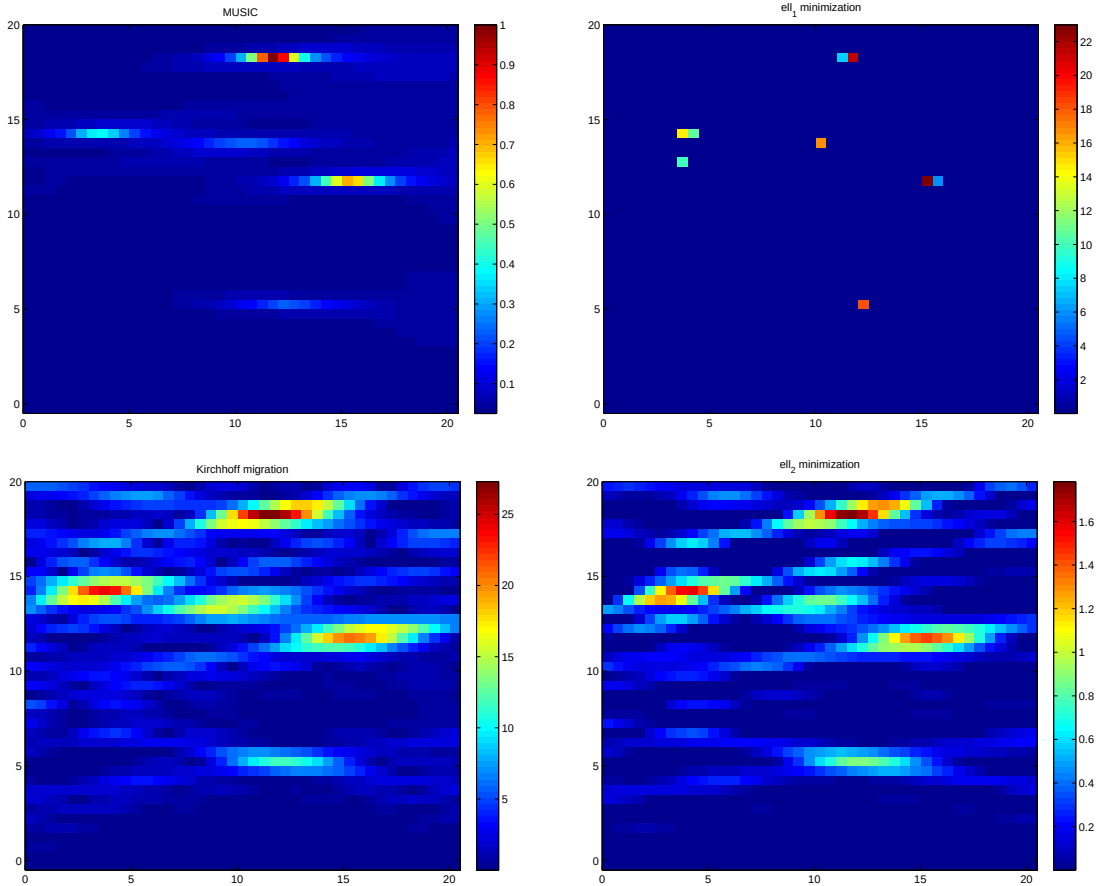


Numerical simulation

MUSIC		ℓ_1		ℓ_2	
loc	val	loc	val	loc	val
$(8, 1, 12)$	1	$(24, 5, 4)$	29.5949	$(8, 1, 12)$	5.5948
$(24, 5, 4)$	0.689	$(31, 9, 17)$	27.6028	$(21, 11, 13)$	4.1529
$(31, 9, 17)$	0.4031	$(8, 1, 12)$	20.4518	$(7, 1, 12)$	3.4482
$(25, 8, 30)$	0.2139	$(21, 11, 13)$	15.4289	$(24, 5, 4)$	3.4439
$(21, 11, 13)$	0.1296	$(25, 8, 30)$	13.4591	$(8, 1, 11)$	3.263
$(21, 10, 13)$	0	$(21, 11, 14)$	0	$(31, 9, 17)$	3.2249
$(31, 10, 17)$	0	$(7, 1, 12)$	0	$(31, 10, 17)$	3.0859
$(25, 9, 30)$	0	$(31, 10, 17)$	0	$(24, 4, 4)$	3.058
$(31, 8, 17)$	0	$(25, 9, 30)$	0	$(24, 6, 4)$	3.0249
$(25, 7, 30)$	0	$(24, 4, 4)$	0	$(31, 8, 17)$	2.6741

Numerical simulation

MUSIC

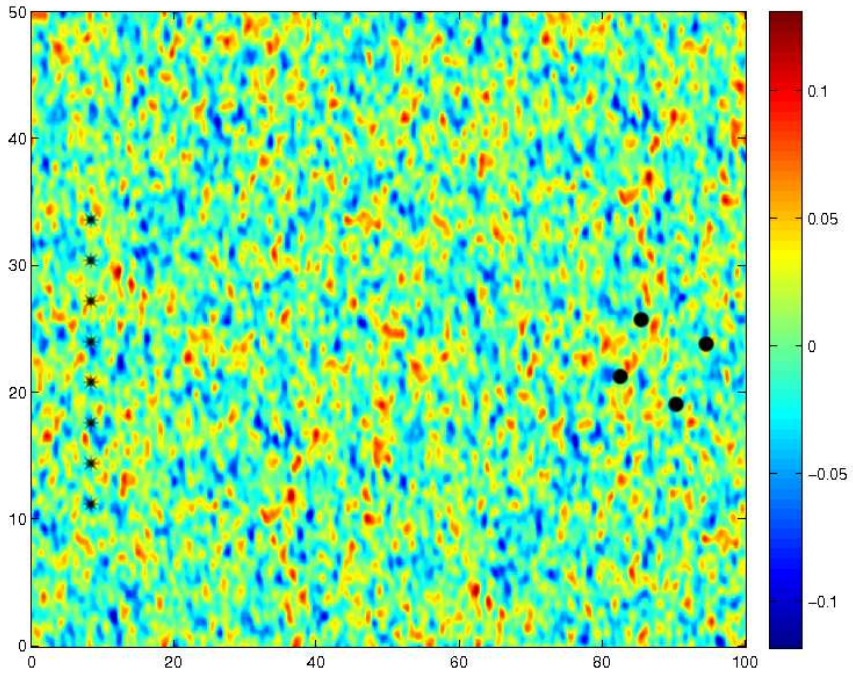


Numerical simulation

MUSIC		ℓ_1		ℓ_2	
loc	val	loc	val	loc	val
$(24, 5, 4)$	1	$(24, 5, 4)$	30.0554	$(8, 1, 12)$	5.8243
$(31, 9, 17)$	0.8352	$(31, 9, 17)$	28.0076	$(21, 11, 13)$	4.3916
$(8, 1, 12)$	0.5154	$(8, 1, 12)$	21.0721	$(7, 1, 12)$	3.5743
$(31, 8, 17)$	0.3179	$(21, 11, 13)$	16.374	$(24, 5, 4)$	3.508
$(31, 10, 17)$	0.3121	$(25, 8, 30)$	14.5407	$(8, 1, 11)$	3.3794
$(24, 6, 4)$	0.3078	$(31, 8, 17)$	0.1677	$(31, 9, 17)$	3.2939
$(24, 4, 4)$	0.2941	$(21, 10, 13)$	0.139	$(31, 10, 17)$	3.1467
$(21, 11, 13)$	0.2711	$(25, 7, 30)$	0.1294	$(24, 4, 4)$	3.1138
$(8, 2, 12)$	0.2375	$(24, 6, 4)$	0.0972	$(24, 6, 4)$	3.0835
$(25, 8, 30)$	0.2225	$(8, 2, 12)$	0.0874	$(21, 11, 14)$	2.7385

Random medium

Phase only perturbation model of random medium.



Realization of random medium with strength of fluctuation $\sigma_0 = 3\%$ and correlation length half of wavelength

Formulation: random medium

Wave speed in a random medium is modeled by

$$c^{-2} = c_0^{-2}(1 + \sigma_0 \mu)$$

where $\mu(\cdot)$ is zero mean, isotropic, weakly stationary random fluctuation. Green function in random medium is modeled by

$$\widehat{G}(\mathbf{x}, \mathbf{y}, \omega) = \widehat{G}_0(\mathbf{x}, \mathbf{y}, \omega) e^{i \frac{\omega \sigma_0}{c_0} |\mathbf{x}-\mathbf{y}| \int_0^1 \mu\left(\frac{\mathbf{x}}{\ell} + \frac{s}{\ell}(\mathbf{y}-\mathbf{x})\right) \mathrm{d}s}$$

and response matrix is obtained by substituting $\widehat{G}_0$ by $\widehat{G}$. As in the case of additive noise, inequality constraints are used in **LSQ** and **CS**, but the noise level ε is usually impossible to estimate accurately, which limits the use of minimization approaches in random medium.

Analysis: moment formula & effective aperture

The moment formula holds for random medium model:

$$\mathbb{E}\left(\widehat{G}(\mathbf{x},\mathbf{y},\omega)\overline{\widehat{G}(\mathbf{x},\mathbf{y}',\omega)}\right) \approx \widehat{G}_0(\mathbf{x},\mathbf{y},\omega)\overline{\widehat{G}_0(\mathbf{x},\mathbf{y}',\omega)}e^{-\frac{\left(\frac{\omega}{c_0}\right)^2|\mathbf{y}-\mathbf{y}'|^2}{2L^2}a_e^2},$$

where

$$a_e = \sigma_0 L\left(-R(0) - \frac{2}{3}\frac{L}{\ell}\int_0^\infty \mathrm{d}t\frac{R'(t)}{t}\right)^{1/2}$$

is the effective aperture of random medium.

Analysis: statistical stability

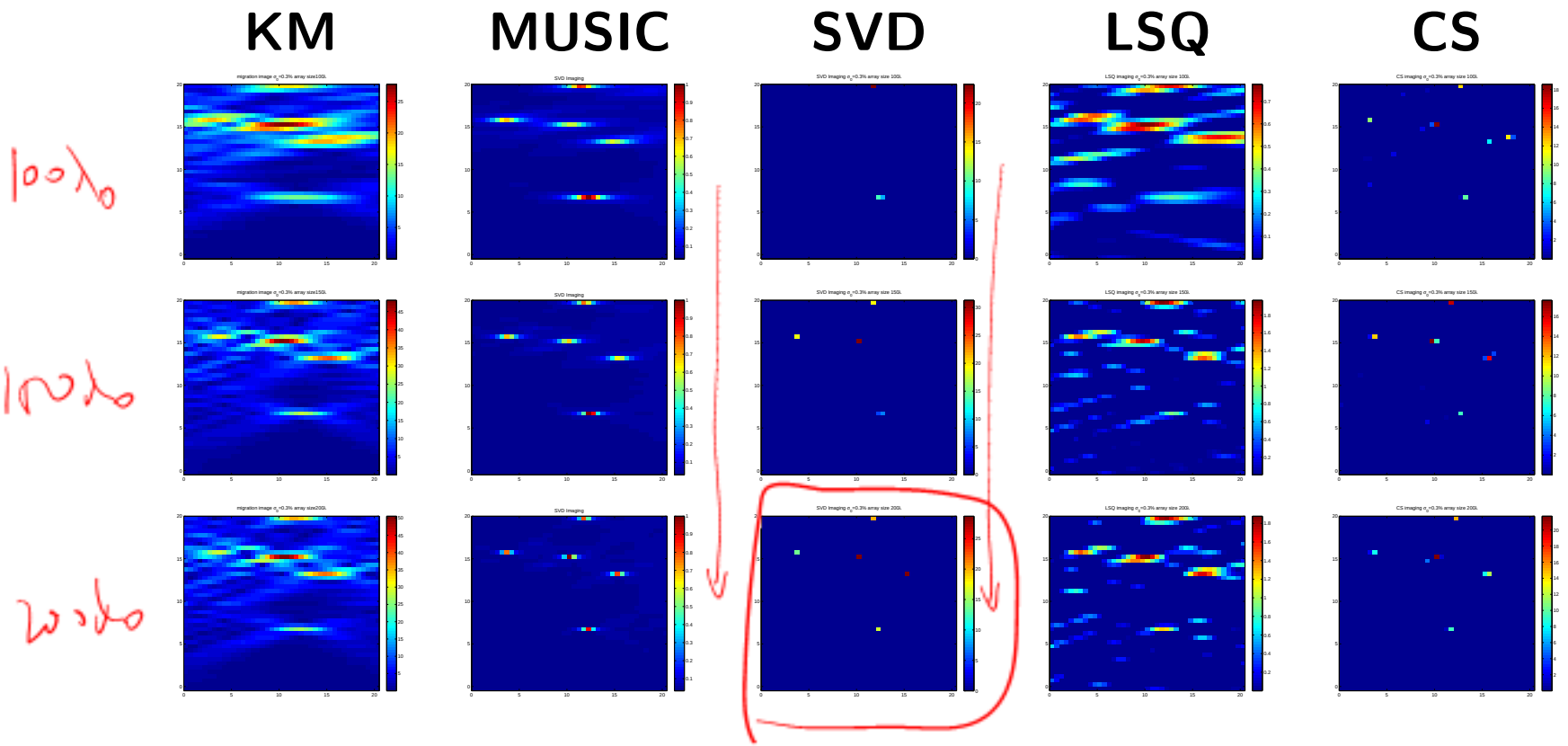
As size of array becomes larger, the random medium model satisfies

$$\mathbb{E}\left|\widehat{G}(\mathbf{x},\mathbf{y},\omega)\overline{\widehat{G}(\mathbf{x},\mathbf{y}',\omega)}-\mathbb{E}\left(\widehat{G}(\mathbf{x},\mathbf{y},\omega)\overline{\widehat{G}(\mathbf{x},\mathbf{y}',\omega)}\right)\right|^2$$

$$\approx \frac{1}{(4\pi)^4|\mathbf{x}-\mathbf{y}|^2|\mathbf{x}-\mathbf{y}'|^2}\left(1-e^{-\left(\frac{\omega}{c_0}\right)^2\frac{a_e^2}{L^2}|\mathbf{y}-\mathbf{y}'|^2}\right),$$

$$\frac{\mathbb{E}\left|\widehat{g}^H(\mathbf{y},\omega)\widehat{g}(\mathbf{y}',\omega)-\mathbb{E}\left(\widehat{g}^H(\mathbf{y},\omega)\widehat{g}(\mathbf{y}',\omega)\right)\right|^2}{\mathbb{E}\left|\widehat{g}(\mathbf{y},\omega)\right|^2}\to 0, \qquad a\to\infty.$$

Numerical simulation – random medium



Array length: $100\lambda_0$ (top), $150\lambda_0$ (middle), $200\lambda_0$ (bottom). Fluctuation of random medium is 0.3%. Resolution improvement is clear for **KM, MUSIC, SVD** which implies statistical stability

Conclusion: Performance of methods

KM: Simple to implement for large arrays, low resolution, wide area imaging, robust to additive noise and random media

LSQ: More costly to implement for large arrays, improves resolution, localized or wide area imaging, not so robust to noise and random media

MUSIC: Simple to implement for small (sparse) targets using the SVD, has good resolution, relatively robust to noise and random media; the modified method can be the basis of more elaborate imaging algorithms to achieve various effects

CS: Very effective for sparse targets in homogeneous media, fast algorithm, excellent accuracy even with substantial subsampling. Not robust to noise or random media (physical experiment by group from Duke).