Stability of Modified-CS over Time for recursive causal sparse reconstruction

Namrata Vaswani

Department of Electrical and Computer Engineering
Iowa State University
http://www.ece.iastate.edu/∼namrata

Recursive Causal Sparse Reconstruction

- Causally & recursively recons. a time seq. of sparse signals
- with slowly changing sparsity patterns
- from as few linear measurements at each time as possible
  - "recursive": use current measurements & previous reconstruction to get current reconstruction

- Potential applications
  - real-time dynamic MRI, e.g. for interventional radiology apps
  - single-pixel video imaging with a real-time video display, ...
  - need: (a) fast acquisition (fewer measurements); (b) process w/o buffering (causal); (c) fast reconstruction (recursive)

Recursive Causal Sparse Reconstruction

- Causally & recursively recons. a time seq. of sparse signals
- with slowly changing sparsity patterns
- from as few linear measurements at each time as possible
  - "recursive": use current measurements & previous reconstruction to get current reconstruction
- Potential applications
  - real-time dynamic MRI, e.g. for interventional radiology apps
  - single-pixel video imaging with a real-time video display, ...
  - need: (a) fast acquisition (fewer measurements); (b) process w/o buffering (causal); (c) fast reconstruction (recursive)
- Most existing work:
  - is either for static sparse reconstruction or is offline & batch, e.g. [Wakin et al (video)], [Gamper et al,Jan'08 (MRI)], [Jung et al'09 (MRI)]

Notation [Candes,Romberg,Tao'05]

- Notation:
  - $T^c = [1, 2, \dots m] \setminus T$: complement of set T
  - $\|A\|$: induced 2-norm of matrix A
  - A_T: sub-matrix containing columns of A with indices in set T
  - A': denotes the transpose of matrix A

- RIP constant, δ_S: smallest real number s.t. all eigenvalues of ${A_T}' A_T$ lie b/w $1 \pm \delta_S$ whenever $|T| \le S$ [Candes,Romberg,Tao'05]
  - *$\delta_S < 1 \Leftrightarrow$ A satisfies the S-RIP*

- ROP constant, θ_{S_1,S_2}: smallest real number s.t. for disjoint sets, T_1, T_2 with $|T_1| \le S_1, \ |T_2| \le S_2$,

  $$|c_1' {A_{T_1}}' A_{T_2} c_2| \le \theta_{S_1,S_2} \|c_1\|_2 \ \|c_2\|_2$$ [Candes,Romberg,Tao'05]

  - *easy to see:* $\|{A_{T_1}}' A_{T_2}\| \le \theta_{|T_1|,|T_2|}$

Sparse reconstruction

- Reconstruct a sparse signal x, with support N, from $y := Ax$,
  - when $n = \text{length}(y) < m = \text{length}(x)$
- Solved if we can find the sparsest vector satisfying $y = Ax$
  - unique solution if $\delta_{2|N|} < 1$
  - exponential complexity
- Practical approaches (polynomial complexity in m)
  - greedy methods, e.g. MP, OMP,..., CoSaMP [Mallat,Zhang'93], [Pati et al'93],...[Needell,Tropp'08]
  - convex relaxation approaches, e.g. BP, BPDN,..., DS, [Chen,Donoho'95], ..., [Candes,Tao'06],...
- Compressed Sensing (CS) literature [Candes,Romberg,Tao'05], [Donoho'05]
  - provides exact reconstruction conditions and error bounds for the practical approaches

Sparse recon. w/ partly known support [Vaswani,Lu, ISIT'09, IEEE Trans. SP'10]

- Recon a sparse signal, x, with support, N, from $y := Ax$
  - given partial but partly erroneous support "knowledge": T

- Recon a sparse signal, x, with support, N, from $y := Ax$
  - given partial but partly erroneous support "knowledge": T
- Rewrite $N :=$ support(x) as

$$N = T \cup \Delta \setminus \Delta_e$$

  - T: support "knowledge"
  - $\Delta := N \setminus T$: misses in T (unknown)
  - $\Delta_e := T \setminus N_t$: extras in T (unknown)

Sparse recon. w/ partly known support [Vaswani,Lu, ISIT'09, IEEE Trans. SP'10]

- Recon a sparse signal, x, with support, N, from $y := Ax$
  - given partial but partly erroneous support "knowledge": T
- Rewrite $N :=$ support(x) as

$$N = T \cup \Delta \setminus \Delta_e$$

  - T: support "knowledge"
  - $\Delta := N \setminus T$: misses in T (unknown)
  - $\Delta_e := T \setminus N_t$: extras in T (unknown)
- If Δ_e empty: find the signal that is sparsest outside of T

$$\min_{\beta} \|(\beta)_{T^c}\|_0 \ s.t. \ y = A\beta$$

  - if $|\Delta|$ small compared to $|N|$: easier problem

Sparse recon. w/ partly known support [Vaswani,Lu, ISIT'09, IEEE Trans. SP'10]

- Recon a sparse signal, x, with support, N, from $y := Ax$
  - given partial but partly erroneous support "knowledge": T
- Rewrite $N :=$ support(x) as

$$N = T \cup \Delta \setminus \Delta_e$$

  - T: support "knowledge"
  - $\Delta := N \setminus T$: misses in T (unknown)
  - $\Delta_e := T \setminus N_t$: extras in T (unknown)
- If Δ_e empty: find the signal that is sparsest outside of T

$$\min_\beta \|(\beta)_{T^c}\|_0 \ s.t. \ y = A\beta$$

  - if $|\Delta|$ small compared to $|N|$: easier problem
- Same thing also works if Δ_e not empty but small
  - exact recon if $\delta_{|N|+|\Delta_e|+|\Delta|} < 1$

- Modified-CS [Vaswani,Lu, ISIT'09, IEEE Trans. SP, Sept'10]

$$\min_{\beta} \|(\beta)_{T^c}\|_1 \ s.t. \ y = A\beta$$

  - we obtained exact reconstruction conditions
  - exact reconstruction is possible using fewer measurements than CS
    - when misses and extras in T small

- Other related and parallel work:
  - [vonBorries et al, TSP'09, CAMSAP'07]: no exact recon conditions or expts.
  - [Khajenejad et al, ISIT'09]: probabilistic prior on support

Problem formulation

- Measure

$$y_t = Ax_t + w_t, \quad \|w_t\|_2 \leq \epsilon$$

  - $A = H\Phi$, H: measurement matrix, Φ: sparsity basis matrix
  - y_t: measurements $(n \times 1)$
  - x_t: sparsity basis coefficients $(m \times 1)$, $m > n$
  - N_t: support of x_t (set of indices of nonzero elements of x_t)

- Goal: recursively reconstruct x_t from $y_0, y_1, \ldots y_t$,
  - i.e. use only $\hat{x}_{t-1}$ and y_t for reconstructing x_t

Problem formulation

- Measure

$$y_t = Ax_t + w_t, \quad \|w_t\|_2 \le \epsilon$$

  - $A = H\Phi$, H: measurement matrix, Φ: sparsity basis matrix
  - y_t: measurements $(n \times 1)$
  - x_t: sparsity basis coefficients $(m \times 1)$, $m > n$
  - N_t: support of x_t (set of indices of nonzero elements of x_t)

- Goal: recursively reconstruct x_t from $y_0, y_1, \ldots y_t$,
  - i.e. use only $\hat{x}_{t-1}$ and y_t for reconstructing x_t

- Key Assumption:
  - *support of x_t, N_t, changes slowly over time:*

$$|N_t \setminus N_{t-1}| \approx |N_{t-1} \setminus N_t| \ll |N_t|$$

    - empirically verified for dynamic MRI sequences [Lu,Vaswani,ICIP'09]

At $t = 0$: simple CS or modified-CS using prior support knowledge

For $t > 0$,

1. *Modified-CS.* Set $T = \hat{N}_{t-1}$ and compute

$$\hat{x}_{t,modcs} = \arg\min_{\beta} \|(\beta)_{T^c}\|_1 \text{ s.t. } \|y_t - A\beta\|_2 \leq \epsilon$$

Modified-CS for time sequences and noisy measurements

At $t = 0$: simple CS or modified-CS using prior support knowledge

For $t > 0$,

1. *Modified-CS.* Set $T = \hat{N}_{t-1}$ and compute

$$\hat{x}_{t,modcs} = \arg\min_{\beta} \|(\beta)_{T^c}\|_1 \text{ s.t. } \|y_t - A\beta\|_2 \leq \epsilon$$

2. *Estimate Support.* Compute $\tilde{T}$ as

$$\tilde{T} = \{i \in [1, m] : |(\hat{x}_{t,modcs})_i| > \alpha\}$$

3. Output $\hat{x}_{t,modcs}$. Set $\hat{N}_t = \tilde{T}$. Feedback $\hat{N}_t$.

Modified-CS for time sequences and noisy measurements

At $t = 0$: simple CS or modified-CS using prior support knowledge

For $t > 0$,

1. *Modified-CS.* Set $T = \hat{N}_{t-1}$ and compute
$$\hat{x}_{t,modcs} = \arg\min_{\beta} \|(\beta)_{T^c}\|_1 \text{ s.t. } \|y_t - A\beta\|_2 \leq \epsilon$$
2. *Estimate Support.* Compute $\tilde{T}$ as
$$\tilde{T} = \{i \in [1, m] : |(\hat{x}_{t,modcs})_i| > \alpha\}$$
3. Output $\hat{x}_{t,modcs}$. Set $\hat{N}_t = \tilde{T}$. Feedback $\hat{N}_t$.

support errors (initial): $\Delta_t := N_t \setminus T_t$, $\Delta_{e,t} := T_t \setminus N_t$,

support errors (final): $\tilde{\Delta}_t := N_t \setminus \tilde{T}_t$, $\tilde{\Delta}_{e,t} := \tilde{T}_t \setminus N_t$

Key Question: "Stability"

- Easy to bound the reconstruction error at a given time, t
  - result depends on the support errors' sizes $|\Delta_t|$, $|\Delta_{e,t}|$
  - may increase over time

Key Question: "Stability"

- Easy to bound the reconstruction error at a given time, t
  - result depends on the support errors' sizes $|\Delta_t|$, $|\Delta_{e,t}|$
  - may increase over time
- **Key Question: is it "stable"?**
  1. Can we obtain conditions under which time-invariant bounds on $|\Delta_t|$, $|\Delta_{e,t}|$ hold?
     - direct corollary: time-invariant bound on the recon error

Key Question: "Stability"

- Easy to bound the reconstruction error at a given time, t
  - result depends on the support errors' sizes $|\Delta_t|$, $|\Delta_{e,t}|$
  - may increase over time
- **Key Question: is it "stable"?**
  1. Can we obtain conditions under which time-invariant bounds on $|\Delta_t|$, $|\Delta_{e,t}|$ hold?
     - direct corollary: time-invariant bound on the recon error
  2. When are these conditions weaker than those for CS?

Key Question: "Stability"

- Easy to bound the reconstruction error at a given time, t
  - result depends on the support errors' sizes $|\Delta_t|$, $|\Delta_{e,t}|$
  - may increase over time
- **Key Question: is it "stable"?**
  1. Can we obtain conditions under which time-invariant bounds on $|\Delta_t|$, $|\Delta_{e,t}|$ hold?
     - direct corollary: time-invariant bound on the recon error
  2. When are these conditions weaker than those for CS?
  3. When are the bounds small compared to support size?

Existing/parallel work

- Recursive reconstruction of sparse signal sequences
  - simple-CS (CS for each time separately): needs larger n
  - [Cevher et al'08] CS on observ differences (CS-diff): unstable
  - [Angelosant,Giannakis,DSP'09]: assume support does not change w/ time
  - [Vaswani, ICIP'08, IEEE Trans. SP, Aug'10] KF-CS, LS-CS-residual (LS-CS)
- **Except our LS-CS work, none of these show error stability over time**

Existing/parallel work

- Recursive reconstruction of sparse signal sequences
  - simple-CS (CS for each time separately): needs larger n
  - [Cevher et al'08] CS on observ differences (CS-diff): unstable
  - [Angelosant,Giannakis,DSP'09]: assume support does not change w/ time
  - [Vaswani, ICIP'08, IEEE Trans. SP, Aug'10] KF-CS, LS-CS-residual (LS-CS)
- **Except our LS-CS work, none of these show error stability over time**
- Our goals very different from:
  - homotopy methods: speed up optimization but not reduce n
  - reconstruct **one signal** recursively from seq. arriving meas's
  - multiple measurements vector (MMV) problem

- LS-CS stability result [Vaswani,IEEE Trans. SP, Aug'10]
  - is for a signal model with support changes "every-so-often".
  - If the delay b/w support change times is large enough; new coeff.'s increase at least at a certain rate; and n large enough;
  - then "stability" holds.

- LS-CS stability result [Vaswani,IEEE Trans. SP, Aug'10]
  - is for a signal model with support changes "every-so-often".
  - If the delay b/w support change times is large enough; new coeff.'s increase at least at a certain rate; and n large enough;
  - then "stability" holds.
- But, often, e.g. in dynamic MRI, support changes occur at every time

Measurement and Signal Model

$$y_t = Ax_t + w_t, \quad \|w_t\|_2 \leq \epsilon$$

- **Why bounded noise? -**
  - Gaussian noise: error bounds at t hold with "large" probability

Measurement and Signal Model

$$y_t = Ax_t + w_t, \quad \|w_t\|_2 \leq \epsilon$$

- **Why bounded noise? -**
  - Gaussian noise: error bounds at t hold with "large" probability
  - for stability, need the bounds to hold for all $0 \leq t < \infty$
    - will hold w.p. zero

Measurement and Signal Model

$$y_t = Ax_t + w_t, \quad \|w_t\|_2 \leq \epsilon$$

- **Why bounded noise? -**
  - Gaussian noise: error bounds at t hold with "large" probability
  - for stability, need the bounds to hold for all $0 \leq t < \infty$
    - will hold w.p. zero
- **Signal model (model on x_t)**
  - S_a additions and S_a removals from support **at each time**
  - Support size constant at S_0

Measurement and Signal Model

$$y_t = Ax_t + w_t, \quad \|w_t\|_2 \leq \epsilon$$

- **Why bounded noise? -**
  - Gaussian noise: error bounds at t hold with "large" probability
  - for stability, need the bounds to hold for all $0 \leq t < \infty$
    - will hold w.p. zero
- **Signal model (model on x_t)**
  - S_a additions and S_a removals from support **at each time**
  - Support size constant at S_0
  - At all t, there are $2S_a$ coeff's each with mag. $r, 2r, \ldots (d-1)r$
    - and $S_0 - (2d-2)S_a$ elements with mag $M := dr$

Measurement and Signal Model

$$y_t = Ax_t + w_t, \quad \|w_t\|_2 \leq \epsilon$$

- **Why bounded noise?** -
  - Gaussian noise: error bounds at t hold with "large" probability
  - for stability, need the bounds to hold for all $0 \leq t < \infty$
    - will hold w.p. zero
- **Signal model (model on x_t)**
  - S_a additions and S_a removals from support **at each time**
  - Support size constant at S_0
  - At all t, there are $2S_a$ coeff's each with mag. $r, 2r, \ldots (d-1)r$
    - and $S_0 - (2d-2)S_a$ elements with mag $M := dr$
  - At all t, S_a out of $2S_a$ elements at mag. jr increase to $(j+1)r$
    - and the other S_a decrease to $(j-1)r$;
    - $j = 0$: coeff's only increase; $j = d$: coeff's only decrease

Example:

- say $m = 200,\ S_0 = 20, S_a = 2,\ d = 3$
- At any t,
  - there are 4 elements each with magnitude $r, 2r$
    - and (20-8)=12 elements with magnitude $M = 3r$

Example:

- say $m = 200,\ S_0 = 20, S_a = 2,\ d = 3$
- At any t,
  - there are 4 elements each with magnitude $r, 2r$
    - and (20-8)=12 elements with magnitude $M = 3r$
  - any 2 out of the 180 zero elements added to support at mag r

Example:

- say $m = 200,\ S_0 = 20, S_a = 2,\ d = 3$
- At any t,
  - there are 4 elements each with magnitude $r, 2r$
    - and (20-8)=12 elements with magnitude $M = 3r$
  - any 2 out of the 180 zero elements added to support at mag r
  - any 2 out of the 4 with mag r increase to $2r$,
    - the other 2 reduce to zero (removed)

Example:

- say $m = 200,\ S_0 = 20, S_a = 2,\ d = 3$
- At any t,
  - there are 4 elements each with magnitude $r, 2r$
    - and (20-8)=12 elements with magnitude $M = 3r$
  - any 2 out of the 180 zero elements added to support at mag r
  - any 2 out of the 4 with mag r increase to $2r$,
    - the other 2 reduce to zero (removed)
  - any 2 out of the 4 with mag 2r increase to $3r$,
    - the other 2 reduce to r

Obtaining the stability result

Corollary (modified-CS error bound [modification of Jacques,2010])

If $\|w_t\|_2 \leq \epsilon$ *and* $\delta_{|N_t|+|\Delta_t|+|\Delta_{e,t}|} < (\sqrt{2}-1)/2$, *then*

$$\|x_t - \hat{x}_{t,modcs}\|_2 \leq C_1(|N_t| + |\Delta_t| + |\Delta_e|) \leq 8.79\epsilon$$

Obtaining the stability result

Corollary (modified-CS error bound [modification of Jacques,2010])

If $\|w_t\|_2 \leq \epsilon$ *and* $\delta_{|N_t|+|\Delta_t|+|\Delta_{e,t}|} < (\sqrt{2}-1)/2$, *then*

$$\|x_t - \hat{x}_{t,modcs}\|_2 \leq C_1(|N_t| + |\Delta_t| + |\Delta_e|) \leq 8.79\epsilon$$

Simple facts

1. All elements with mag $> b$ definitely detected at t
   - if $b \geq \alpha \ + \ \max_i |(x_t - \hat{x}_{modcs,t})_i|$

Obtaining the stability result

Corollary (modified-CS error bound [modification of Jacques,2010])

If $\|w_t\|_2 \leq \epsilon$ *and* $\delta_{|N_t|+|\Delta_t|+|\Delta_{e,t}|} < (\sqrt{2}-1)/2$, *then*

$$\|x_t - \hat{x}_{t,modcs}\|_2 \leq C_1(|N_t| + |\Delta_t| + |\Delta_e|) \leq 8.79\epsilon$$

Simple facts

1. All elements with mag $> b$ definitely detected at t
   - if $b \geq \alpha \ + \ \max_i |(x_t - \hat{x}_{modcs,t})_i|$
2. All zero elements definitely deleted/not falsely added at t
   - if $\alpha \geq \max_i |(x_t - \hat{x}_{modcs,t})_i|$

Obtaining the stability result

Corollary (modified-CS error bound [modification of Jacques,2010])

If $\|w_t\|_2 \leq \epsilon$ *and* $\delta_{|N_t|+|\Delta_t|+|\Delta_{e,t}|} < (\sqrt{2}-1)/2$, *then*

$$\|x_t - \hat{x}_{t,modcs}\|_2 \leq C_1(|N_t| + |\Delta_t| + |\Delta_e|) \leq 8.79\epsilon$$

Simple facts

1. All elements with mag $> b$ definitely detected at t
   - if $b \geq \alpha \ + \ \max_i |(x_t - \hat{x}_{modcs,t})_i|$
2. All zero elements definitely deleted/not falsely added at t
   - if $\alpha \geq \max_i |(x_t - \hat{x}_{modcs,t})_i|$

- Use above facts/corollary to obtain sufficient conditions s.t.
  - only coeff's with magnitude $< 2r$ are part of missed set, $\tilde{\Delta}_t$,
  - and the final set of extras, $\tilde{\Delta}_{e,t}$ is an empty set

support errors (initial): $\Delta_t := N_t \setminus T_t$, $\Delta_{e,t} := T_t \setminus N_t$, support errors (final): $\tilde{\Delta}_t := N_t \setminus \tilde{T}_t$, $\tilde{\Delta}_{e,t} := \tilde{T}_t \setminus N_t$

Theorem (Stability of Modified-CS)

If

1. *(support estimation threshold)* $\alpha = 8.79\epsilon$
2. *(support size, support change size)* S_0, S_a *satisfy*
   - $\delta_{S_0+3S_a} < (\sqrt{2}-1)/2$ *(for a given A)*
3. *(new coeff. increase rate)* $r \geq 8.79\epsilon$,
4. *(initial time)* *at* $t = 0$, n_0 *large enough s.t.* $\delta_{2S_0} < (\sqrt{2}-1)/2$

then, at all times, t,

Theorem (Stability of Modified-CS)

If

1. *(support estimation threshold)* $\alpha = 8.79\epsilon$
2. *(support size, support change size)* S_0, S_a *satisfy*
   - $\delta_{S_0+3S_a} < (\sqrt{2}-1)/2$ *(for a given A)*
3. *(new coeff. increase rate)* $r \geq 8.79\epsilon$,
4. *(initial time)* *at* $t = 0$, n_0 *large enough s.t.* $\delta_{2S_0} < (\sqrt{2}-1)/2$

then, at all times, t,

- *final support errors,* $|\tilde{\Delta}_t| \leq 2S_a$ *and* $|\tilde{\Delta}_{e,t}| = 0$
- *initial support errors,* $|\Delta_t| \leq 2S_a$ *and* $|\Delta_{e,t}| \leq S_a$

Theorem (Stability of Modified-CS)

If

1. *(support estimation threshold)* $\alpha = 8.79\epsilon$
2. *(support size, support change size)* S_0, S_a *satisfy*
   - $\delta_{S_0+3S_a} < (\sqrt{2}-1)/2$ *(for a given A)*
3. *(new coeff. increase rate)* $r \geq 8.79\epsilon$,
4. *(initial time) at* $t = 0$, n_0 *large enough s.t.* $\delta_{2S_0} < (\sqrt{2}-1)/2$

then, at all times, t,

- *final support errors,* $|\tilde{\Delta}_t| \leq 2S_a$ *and* $|\tilde{\Delta}_{e,t}| = 0$
- *initial support errors,* $|\Delta_t| \leq 2S_a$ *and* $|\Delta_{e,t}| \leq S_a$
- *and so recon error satisfies* $\|x_t - \hat{x}_{t,modcs}\|_2 \leq 8.79\epsilon$

Theorem (Stability of Modified-CS)

If

1. *(support estimation threshold)* $\alpha = 8.79\epsilon$
2. *(support size, support change size)* S_0, S_a *satisfy*
   - $\delta_{S_0+3S_a} < (\sqrt{2}-1)/2$ *(for a given A)*
3. *(new coeff. increase rate)* $r \geq 8.79\epsilon$,
4. *(initial time) at* $t = 0$, n_0 *large enough s.t.* $\delta_{2S_0} < (\sqrt{2}-1)/2$

then, at all times, t,

- *final support errors,* $|\tilde{\Delta}_t| \leq 2S_a$ *and* $|\tilde{\Delta}_{e,t}| = 0$
- *initial support errors,* $|\Delta_t| \leq 2S_a$ *and* $|\Delta_{e,t}| \leq S_a$
- *and so recon error satisfies* $\|x_t - \hat{x}_{t,modcs}\|_2 \leq 8.79\epsilon$

- **Slow support change** $\Rightarrow S_a \ll S_0$
  - $\Rightarrow$ **support errors' bound small compared to support size**

Discussion

Compare with simple CS

- To get the same error bound, CS needs
  - $\delta_{2S_0} < (\sqrt{2} - 1)/2$
- Modified-CS only needs
  - $\delta_{S_0+3S_a} < (\sqrt{2} - 1)/2$
    - recall: S_0: support size, S_a: # of support changes at t

Discussion

Compare with simple CS

- To get the same error bound, CS needs
  - $\delta_{2S_0} < (\sqrt{2} - 1)/2$
- Modified-CS only needs
  - $\delta_{S_0+3S_a} < (\sqrt{2} - 1)/2$
    - recall: S_0: support size, S_a: # of support changes at t

Limitations

- Bounding ℓ_∞ norm of error by ℓ_2 norm: loose

Discussion

Compare with simple CS

- To get the same error bound, CS needs
  - $\delta_{2S_0} < (\sqrt{2} - 1)/2$
- Modified-CS only needs
  - $\delta_{S_0+3S_a} < (\sqrt{2} - 1)/2$
    - recall: S_0: support size, S_a: # of support changes at t

Limitations

- Bounding ℓ_∞ norm of error by ℓ_2 norm: loose
- Using a single threshold, α, for simultaneous add/del to/from support
  - need α large enough to ensure correct deletion
  - $\Rightarrow$ need rate of coeff. increase, r, even larger

A two threshold solution: Add-LS-Del [1]

- Add using a small threshold

$$T_{\text{add}} = T \cup \{i : |(\hat{x}_{\text{modCS}})_i| > \alpha_{\text{add}}\}$$

  - can use α_{add} just large enough s.t. well-conditioned $(A)_{T_{\text{add}}}$

[1] idea related to [DantzigSelector,Candes,Tao'06], [KF-CS,Vaswani'08], [CoSaMP,Needell,Tropp'08]

A two threshold solution: Add-LS-Del [1]

- Add using a small threshold

$$T_{\text{add}} = T \cup \{i : |(\hat{x}_{\text{modCS}})_i| > \alpha_{\text{add}}\}$$

  - can use α_{add} just large enough s.t. well-conditioned $(A)_{T_{\text{add}}}$

- Compute LS estimate on T_{add}

$$\hat{x}_{\text{add}} = LS(T_{\text{add}}, y_t)$$

  - reduces bias and mean squared error if $T_{\text{add}} \approx N_t$

[1] idea related to [DantzigSelector,Candes,Tao'06], [KF-CS,Vaswani'08], [CoSaMP,Needell,Tropp'08]

A two threshold solution: Add-LS-Del [1]

- Add using a small threshold

$$T_{\text{add}} = T \cup \{i : |(\hat{x}_{\text{modCS}})_i| > \alpha_{\text{add}}\}$$

  - can use α_{add} just large enough s.t. well-conditioned $(A)_{T_{\text{add}}}$

- Compute LS estimate on T_{add}

$$\hat{x}_{\text{add}} = LS(T_{\text{add}}, y_t)$$

  - reduces bias and mean squared error if $T_{\text{add}} \approx N_t$

- Delete with larger threshold

$$\hat{N} = T_{add} \setminus \{i : |(\hat{x}_{add})_i| \le \alpha_{\text{del}}\}$$

  - only deleting (not adding) $\Rightarrow \alpha_{\text{del}}$ can be larger
  - $\hat{x}_{\text{add}}$ more accurate $\Rightarrow \alpha_{\text{del}}$ can be larger

[1] idea related to [DantzigSelector,Candes,Tao'06], [KF-CS,Vaswani'08], [CoSaMP,Needell,Tropp'08]

Obtaining the stability result

Lemma (Detection condition)

All elements with magnitude $> b$ definitely detected at t if

- $\|w\| \leq \epsilon$, $\delta_{S_0+|\Delta_{e,t}|+|\Delta_t|} < (\sqrt{2}-1)/2$ and $b > \alpha_{add} + 8.79\epsilon$

Lemma (No false deletion condition)

All elements in T_{add} with magnitude $> b$ not deleted at t if

- $\|w\| \leq \epsilon$, $\delta_{|T_{add}|} < 1/2$ and $b_1 > \alpha_{del} + \sqrt{2}\epsilon + 2\theta_{|T_{add}|,|\Delta_{add}|}\|x_{\Delta_{add}}\|_2$

Lemma (Deletion condition)

All elements of $\Delta_{e,add,t}$ deleted at t if

- $\|w\| \leq \epsilon$, $\delta_{|T_{add}|} < 1/2$ and $\alpha_{del} \geq \sqrt{2}\epsilon + 2\theta_{|T_{add}|,|\Delta_{add}|}\|x_{\Delta_{add}}\|_2$

$$\text{From the signal model, } N_t = N_{t-1} \cup \mathcal{A}_t \setminus \mathcal{R}_t$$
$$\mathcal{S}_{t,2} = \mathcal{S}_{t-1,2} \cup (\mathcal{A}_t \cup \mathcal{D}_{t,1}) \setminus (\mathcal{R}_t \cup \mathcal{I}_{t,2})$$

$\mathcal{S}_{t,2}$: set of indices of all nonzero coeff's with magnitude $< 2r$

$\mathcal{A}_t$: new additions at t, $\mathcal{R}_t$: new removals at t

$\mathcal{I}_{t,2}$: all coeff's that increased from r to $2r$ at t, $\mathcal{D}_{t,1}$: decreased from $2r$ to r

Theorem (Stability of modified-CS with add-LS-del)

If

1. *(addition and deletion thresholds)*
   - α_{add} *is large enough s.t. at most* S_a *false adds per unit time,*
   - $\alpha_{del} = \sqrt{2}\epsilon + 2\sqrt{S_a}\theta_{S_0+2S_a,S_a} r$,
2. *(support size, support change size)* S_0, S_a *satisfy*
   - $\delta_{S_0+3S_a} < (\sqrt{2}-1)/2$, *and*
   - $\theta_{S_0+2S_a,S_a} < \frac{1}{4\sqrt{S_a}}$,
3. *(new coeff. increase rate)* $r \geq \max(G_1, G_2)$, *where*

$$G_1 \triangleq \frac{\alpha_{add} + 8.79\epsilon}{2}, \quad G_2 \triangleq \frac{\sqrt{2}\epsilon}{1 - 2\sqrt{S_a}\theta_{S_0+2S_a,S_a}}$$

4. *(initial time)* *at* $t = 0$, n_0 *is large enough*

then, at all t, *all the same conclusions hold.*

Discussion – 1: Limitation and a Way Out

- $\theta_{S_0+2S_a,S_a} < 1/(4\sqrt{S_a})$ difficult to satisfy for large problems

Discussion – 1: Limitation and a Way Out

- $\theta_{S_0+2S_a,S_a} < 1/(4\sqrt{S_a})$ difficult to satisfy for large problems
- Get this since we bound LS error as $\|x - \hat{x}_{\text{add}}\|_\infty \le \|x - \hat{x}_{\text{add}}\|_2$
  - clearly a loose bound
  - esp. since LS step reduces bias (when support errors small)

Discussion – 1: Limitation and a Way Out

- $\theta_{S_0+2S_a,S_a} < 1/(4\sqrt{S_a})$ difficult to satisfy for large problems
- Get this since we bound LS error as $\|x - \hat{x}_{\text{add}}\|_\infty \leq \|x - \hat{x}_{\text{add}}\|_2$
  - clearly a loose bound
  - esp. since LS step reduces bias (when support errors small)
- Instead if assume $\|x - \hat{x}_{\text{add}}\|_\infty \leq (1/\sqrt{S_a})\ \|x - \hat{x}_{\text{add}}\|_2$, then
  - **theta condition weakened to**

$$\theta_{S_0+2S_a,S_a} < 1/4$$

  - and lower bound on coeff. increase rate, r, also reduced

Discussion – 1: Limitation and a Way Out

- $\theta_{S_0+2S_a,S_a} < 1/(4\sqrt{S_a})$ difficult to satisfy for large problems
- Get this since we bound LS error as $\|x - \hat{x}_{\text{add}}\|_\infty \leq \|x - \hat{x}_{\text{add}}\|_2$
  - clearly a loose bound
  - esp. since LS step reduces bias (when support errors small)
- Instead if assume $\|x - \hat{x}_{\text{add}}\|_\infty \leq (1/\sqrt{S_a})\ \|x - \hat{x}_{\text{add}}\|_2$, then
  - **theta condition weakened to**

$$\theta_{S_0+2S_a,S_a} < 1/4$$

  - and lower bound on coeff. increase rate, r, also reduced
  - (in simulation expts, above assumption holds 99% of times)

Discussion – 2: Comparisons

Comparison with CS result

- For the same error bound, CS needs:

$$\delta_{2S_0} < (\sqrt{2} - 1)/2$$

- Mod-CS with add-LS-del only needs:

$$\delta_{S_0+3S_a} < (\sqrt{2} - 1)/2 \text{ and } \theta_{S_0+2S_a,S_a} < 1/4$$

Comparison with Modified-CS result

- Mod-CS needs $r \geq 8.79\epsilon$
- Mod-CS with add-LS-del only needs $r \geq (\alpha_{\text{add}} + 8.79\epsilon)/2$
  - usually α_{add} can be quite small

Discussion – 2: Comparisons

Comparison with CS result

- For the same error bound, CS needs:

$$\delta_{2S_0} < (\sqrt{2} - 1)/2$$

- Mod-CS with add-LS-del only needs:

$$\delta_{S_0+3S_a} < (\sqrt{2} - 1)/2 \text{ and } \theta_{S_0+2S_a,S_a} < 1/4$$

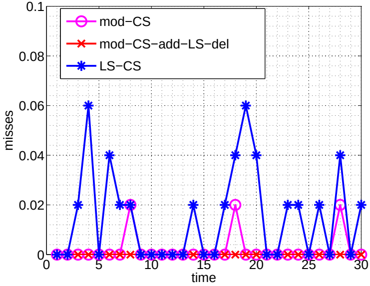
Comparison with Modified-CS result

- Mod-CS needs $r \geq 8.79\epsilon$
- Mod-CS with add-LS-del only needs $r \geq (\alpha_{\text{add}} + 8.79\epsilon)/2$
  - usually α_{add} can be quite small

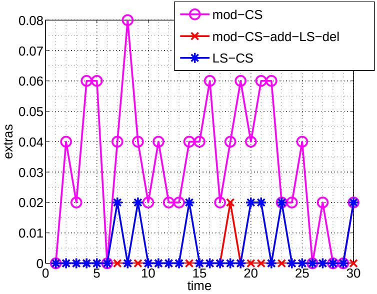
Comparison with LS-CS result

- proved similar result for LS-CS; its requirements much stronger

Simulations: support errors

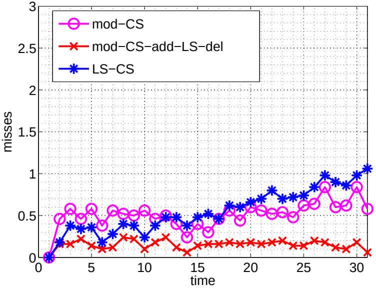


(a) $r = 1$: (mean # of misses)$/S_0$

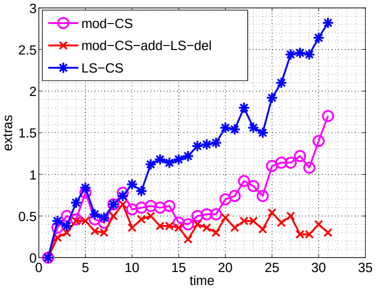


(b) $r = 1$: (mean # of extras)$/S_0$

- Measurement model: $n = 29.5\%$, $w_t \sim unif(-c, c)$ with $c = 0.1266$
- Support size, $S_0 = 10\%$, support change size, $S_a = 1\%$
- Signal model: $r = 1$, $d = 3$



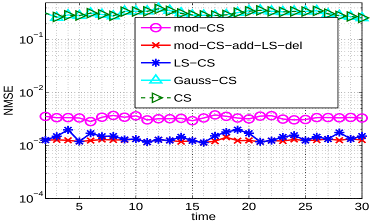
(c) $r = 1/2$: (mean # of misses)$/S_0$



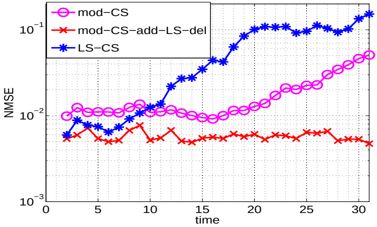
(d) $r = 1/2$: (mean # of extras)$/S_0$

- Measurement model: $n = 29.5\%$, $w_t \sim unif(-c, c)$ with $c = 0.1266$
- Support size, $S_0 = 10\%$, support change size, $S_a = 1\%$
- Signal model: $r = 1/2$, $d = 4$

Simulations: reconstruction error



$r = 1$



$r = 1/2$

Conclusions and Ongoing Work

- Under mild assumptions (S_0, S_a small enough and r large enough), we obtained time-invariant support error (and recon. error) bounds for
  - modified-CS (single threshold)
  - modified-CS with add-LS-del
- If "slow support change" holds, i.e. if $S_a \ll S_0$,
  - the support error bounds are small compared to support size
  - larger support size is allowed than what simple CS needs

Conclusions and Ongoing Work

- Under mild assumptions (S_0, S_a small enough and r large enough), we obtained time-invariant support error (and recon. error) bounds for
  - modified-CS (single threshold)
  - modified-CS with add-LS-del
- If "slow support change" holds, i.e. if $S_a \ll S_0$,
  - the support error bounds are small compared to support size
  - larger support size is allowed than what simple CS needs
- Ongoing work
  - Experiments with real functional MRI sequences
  - Stability of KalMoCS (Kalman-like Modified-CS)
    - Mod-CS with a slow signal value change term
  - Real-time (recursive and causal) robust PCA [Qiu,Vaswani, Allerton'10]
    - online matrix completion w/ sparse corruptions

Modified-CS stability (English version)

For a given measurement matrix, A, and noise bound, ϵ, if,

1. the support estimation threshold(s) are large enough,
2. the support size, S_0, and support change size, S_a are small enough,
3. the newly added coefficients increase (existing large coefficients decrease) at least at a certain rate, r, and
4. the initial number of measurements, n_0, is large enough for simple CS

then

Modified-CS stability (English version)

For a given measurement matrix, A, and noise bound, ϵ, if,

1. the support estimation threshold(s) are large enough,
2. the support size, S_0, and support change size, S_a are small enough,
3. the newly added coefficients increase (existing large coefficients decrease) at least at a certain rate, r, and
4. the initial number of measurements, n_0, is large enough for simple CS

then

- the support errors are bounded by time-invariant values
  - $|N_t \setminus \hat{N}_{t-1}| \leq 2S_a$, $|\hat{N}_{t-1} \setminus N_t| \leq S_a$

Modified-CS stability (English version)

For a given measurement matrix, A, and noise bound, ϵ, if,

1. the support estimation threshold(s) are large enough,
2. the support size, S_0, and support change size, S_a are small enough,
3. the newly added coefficients increase (existing large coefficients decrease) at least at a certain rate, r, and
4. the initial number of measurements, n_0, is large enough for simple CS

then

- the support errors are bounded by time-invariant values
  - $|N_t \setminus \hat{N}_{t-1}| \leq 2S_a$, $|\hat{N}_{t-1} \setminus N_t| \leq S_a$
- consequently, the recon. error is also "stable"
- "Slow support change" $\Rightarrow S_a \ll S_0 \Rightarrow$ support error bound small

Proof Outline: Proof by induction

To show: under Theorem 1 conditions, $|\tilde{\Delta}_{e,t}| = 0$; $\tilde{\Delta}_t \subseteq \mathcal{S}_{t,2}$

1. bound $|\Delta_t|$, $|\Delta_{e,t}|$, $|T_t|$
   - by induc. assump., $|T_t| = |\tilde{T}_{t-1}| \leq |N_{t-1}| + |\tilde{\Delta}_{e,t-1}| \leq S_0$
   - use signal model & induc. assump. to bound $|\Delta_t|$, $|\Delta_{e,t}|$
2. bound $|\Delta_{\text{add},t}|$, $|\Delta_{\text{add},e,t}|$, $|T_{\text{add},t}|$
   - use 1; detection conditions; and following[2] to bound $\Delta_{\text{add},t}$
     $$\mathcal{S}_{t,2} = \mathcal{S}_{t-1,2} \cup (\mathcal{A}_t \cup \mathcal{D}_{t,1}) \setminus (\mathcal{R}_t \cup \mathcal{I}_{t,2})$$
   - use 1 and bound on # of false adds to show $|\Delta_{e,\text{add},t}| \leq 2S_a$; and so $|T_{\text{add},t}| \leq |N_t| + 2S_a = S_0 + 2S_a$
3. bound $|\tilde{\Delta}_t|$, $|\tilde{\Delta}_{e,t}|$
   - use 2 and no-false-deletion conditions to show $\tilde{\Delta}_t \subseteq \mathcal{S}_{t,2}$
   - use deletion condition lemma to show $|\tilde{\Delta}_{e,t}| = 0$

---

[2] $\mathcal{S}_{t,2}$: set of indices of all nonzero coeff's with magnitude $< 2r$

$\mathcal{I}_{t,2}$: all coeff's that increased from r to $2r$ at t, $\mathcal{D}_{t,1}$: decreased from $2r$ to r

$\mathcal{A}_t$: new additions at t, $\mathcal{R}_t$: new removals at t