KF-CS Theory

Animesh Biswas, Kevin Palmowski

Last compiled June 12, 2014

1 Model

At each time t, we have $y_t = Ax_t + w_t$.

Time indices are discrete. Make the distinction between sampling times (used) and continuous time (not used).

2 Algorithm – KF-CS with LS

3 Proofs

Lemma 1. *Assume that $\{x_t\}$ follow the signal model above, $y_t = Ax_t + w_t$, $\{t_0, t_0+1, t_0+2, \ldots\}$ is a discrete set of sampling times, only additions to true support ($N_t \subseteq N_{t+1}$ for all t), etc.*

Further assume that

i) The true solution is exactly recovered at the initial time t_0: $\hat{x}_{t_0} = x_{t_0}$, so $\hat{N}_{t_0} = N_{t_0} = N_0$; ***Can we relax this to just the true support is recovered?***

ii) The maximum support size S_{max} satisfies $S_{max} \leq S_{**} = \max\{s : \delta_{2s}(A) < \sqrt{2}-1\}$;

iii) The observation noise w_t is bounded in magnitude: $\|w_t\| < \xi$ for all t and some $\xi > 0$;

iv) The addition threshold α_t satisfies $\alpha_t = C_1\xi$ for each sampling time t, where $C_1 = C_1(|N_t|, \xi)$ ***(verify)*** *is defined* ***below OR in Candes****; and*

Note: if we can define α as max of all $C_1\xi$, we can remove t dependence. Alternately since $|N_t| \leq S_{max}$ for all t, we can define C_1 in terms of S_{max} and if C_1 is increasing we can remove the t dependence that way.

*v) The addition delay d satisfies $d > \tau_{det}$ (****PROBLEM: τ_{det} is a function of t, which destroys the entire argument – need to remove t dependence in α****), where the detection delay τ_{det} is defined by*

$$\tau_{det} = \tau_{det}(\alpha_t, \varepsilon) = \left\lceil \left(\frac{2\alpha_t}{\sigma_{sys}\mathbf{\Phi}^{-1}\left(\frac{(1-\varepsilon)^{1/S_a}}{2}\right)} \right)^2 \right\rceil . \leftarrow \textbf{\textit{superscript looks bad}}$$

Here, $\mathbf{\Phi}^{-1}(x)$ is the inverse of the standard Guassian CDF, $\mathbf{\Phi}(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dz$.

Then

1) $\|x_t - \hat{x}_{t,CSres}\|_2 \leq \alpha_t$ for all sampling times t;

2) $\hat{N}_t \subseteq N_t$ for all sampling times t; and

3) $\boldsymbol{Pr}(E_j \mid F_j) \geq 1-\varepsilon$, where $E_j = \{\hat{N}_t = N_t \text{ for all } t \in [t_j + \tau_{det} : t_{j+1}-1]\}$ and $F_j = \{\hat{N}_{t_j - 1} = N_{t_j - 1}\}$.

Proof. **Need to find some way to get Candes Thm 1.3 in here and make the connection that $\hat{x}_{t,\mathbf{CSres}}$ in our notation is $x^\star$ in his**

To prove claims 1 and 2, we proceed by induction on the value of t.

Consider the base case, where $t = t_0$. Claim 1 follows from Theorem 1.3 in [1] and assumptions (ii), (iii), and (iv) **(Not immediate – need to connect to Candes as above)**, and assumption (i) trivially proves claim 2.

Suppose now that claims 1 and 2 are both true for some time $(t-1)$. We show that the claims are true at time t.

First, we verify claim 1 at time t. Let

$$\begin{aligned}
\beta_t &= x_t - \hat{x}_{t,\text{init}} \\
\hat{\beta}_t &= \arg\min_{\beta} \|\beta\|_1 \text{ subject to } \|y_t - A\beta\|_2 < \xi \\
\hat{x}_{t,\text{CSres}} &= \hat{x}_{t,\text{init}} + \hat{\beta}_t,
\end{aligned}$$

where $\hat{x}_{t,\text{init}}$ is defined in the algorithm and $\text{supp}(\hat{x}_{t,\text{init}}) = \hat{N}_{t-1}$.

By the induction hypothesis, $\hat{N}_{t-1} \subseteq N_{t-1}$, and by our model assumptions we have $N_{t-1} \subseteq N_t$. Therefore, $\text{supp}(\beta_t) \subseteq N_t \cup N_{t-1} = N_t$, so $|\,\text{supp}(\beta_t)| \leq |N_t| \leq S_{\max}$. With this, we can apply Theorem 1.3 in [1] to see that $\|\beta_t - \hat{\beta}_t\|_2 \leq \alpha_t$ **(AGAIN, need to make this connection)**. By the definitions of β_t and $\hat{x}_{t,\text{CSres}}$, we see that $\|\beta_t - \hat{\beta}_t\|_2 = \|x_t - \hat{x}_{t,\text{CSres}}\|_2$, so claim 1 follows.

Next, we verify claim 2 at time t. Suppose that $(x_t)_i = 0$ for some index i, so that $i \notin \text{supp}(x_t) = N_t$. Since $N_{t-1} \subseteq N_t$, we must also have $i \notin N_{t-1}$; by the induction hypothesis, this implies that $i \notin \hat{N}_{t-1}$.

Applying the result of claim 1,

$$|(\hat{x}_{t,\text{CSres}})_i|^2 = |(x_t - \hat{x}_{t,\text{CSres}})_i|^2 \leq \|x_t - \hat{x}_{t,\text{CSres}}\|_2^2 \leq \alpha_t^2,$$

so $|(\hat{x}_{t,\text{CSres}})_i| \leq \alpha_t$. Referring to the algorithm, $\hat{N}_t = \hat{N}_{t-1} \cup \{j : |(\hat{x}_{t,\text{CSres}})_j| > \alpha_t\}$. Since $i \notin \hat{N}_{t-1}$ and $|(\hat{x}_{t,\text{CSres}})_i| \leq \alpha_t$, it follows that $i \notin \hat{N}_t$. Thus if $i \notin N_t$, then $i \notin \hat{N}_t$; equivalently, if $i \in \hat{N}_t$, then $i \in N_t$. Therefore, $\hat{N}_t \subseteq N_t$, which proves claim 2 and completes our induction proof.

Now, we prove claim 3. Let $t \in [t_j : t_{j+1} - 1]$ for some $j \geq 0$.

Let $\Delta_t = N_t \setminus \hat{N}_{t-1}$ denote the set of indices of the true support at time t which have not been detected before time t.

Let $i \in \Delta_t$.

Suppose $|(x_t)_i| > 2\alpha_t$, which implies that

$$\begin{aligned}
|(\hat{x}_{t,\text{CSres}})_i| &= |(x_t)_i - [(x_t)_i - (\hat{x}_{t,\text{CSres}})_i]| \\
&\geq \big||(x_t)_i| - |(x_t - \hat{x}_{t,\text{CSres}})_i|\big| \\
&\geq \big||(x_t)_i| - \|(x_t - \hat{x}_{t,\text{CSres}})_i\|_2\big| \\
&> |2\alpha_t - \alpha_t| \\
&= \alpha_t,
\end{aligned}$$

where have applied the result from claim 1 in the final inequality. We see that if $|(x_t)_i| > 2\alpha_t$, then $|(\hat{x}_{t,\text{CSres}})_i| > \alpha_t$, so $i \in \hat{N}_t = \hat{N}_{t-1} \cup \{j : |(\hat{x}_{t,\text{CSres}})_j| > \alpha_t\}$.

So if $|(x_t)_j| > 2\alpha_t$ for all $j \in \Delta_t$, then $\Delta_t \subseteq \{j : |(\hat{x}_{t,\text{CSres}})_j| > \alpha_t\} \subseteq \hat{N}_t$, that is, each undetected element of the true support will be detected at time t.

The event $\{\hat{N}_t = N_t \mid F_j\}$ is equivalent to the event $\{\Delta_t \subseteq \hat{N}_t\}$.

Conditioned on F_j, we have $N_{t_j} = N_{t_j - 1} \cup \Delta_{\text{add},t_j} = \hat{N}_{t_j - 1} \cup \Delta_{\text{add},t_j}$. So we add at most S_{add} new indices.

FIX ALL OF the stuff above this line, it's a disaster... moving on and skipping some stuff...

The entries $(x_t)_i$ are independent and identically distributed $\mathcal{N}\left(0, (t - t_j)\sigma_{\text{sys}}^2\right)$ random variables.

We see that **(we can get rid of some of the F_j conditions somewhere in here)**

$$\begin{aligned}
\mathbf{Pr}(\hat{N}_t = N_t \mid F_j) &\geq \mathbf{Pr}\left(|(x_t)_i| > 2\alpha_t \text{ for all } i \in \Delta_t \mid F_j\right) \\
&\geq \mathbf{Pr}\left(|(x_t)_i| > 2\alpha_t \text{ for all } i \in \Delta_{\text{add},t_j} \mid F_j\right) \\
&= \left[\mathbf{Pr}\left(|(x_t)_i| > 2\alpha_t \text{ for some } i \in \Delta_{\text{add},t_j}\right)\right]^{|\Delta_{\text{add},t_j}|} \\
&= \left[\mathbf{Pr}\left(|(x_t)_1| > 2\alpha_t\right)\right]^{S_{\text{add}}} \\
&= \left[2\boldsymbol{\Phi}\left(\frac{2\alpha_t}{\sigma_{\text{sys}}\sqrt{t - t_j}}\right)\right]^{S_{\text{add}}}.
\end{aligned}$$

We examine the particular case there $t = t_j + \tau_{\mathrm{det}}$. In this case,

$$\begin{aligned}
\mathbf{Pr}(\hat{N}_{t_j+\tau_{\mathrm{det}}} = N_{t_j+\tau_{\mathrm{det}}} \mid F_j) &\geq \left[2\mathbf{\Phi}\left(\frac{2\alpha_t}{\sigma_{\mathrm{sys}}\sqrt{(t_j+\tau_{\mathrm{det}}) - t_j}}\right)\right]^{S_{\mathrm{add}}} \\
&= \left[2\mathbf{\Phi}\left(\frac{2\alpha_t}{\sigma_{\mathrm{sys}}\sqrt{\tau_{\mathrm{det}}}}\right)\right]^{S_{\mathrm{add}}} \\
&\textbf{FILL IN THE DETAILS after we get } \boldsymbol{\tau_{\mathrm{det}}} \textbf{ settled...} \\
&\geq 1 - \varepsilon.
\end{aligned}$$

If $\hat{N}_t = N_t$ for $t = t_j + \tau_{\mathrm{det}}$, then the model and algorithm assumptions imply that $\hat{N}_t = N_t$ for all $t \in [t_j + \tau_{\mathrm{det}} : t_{j+1} - 1]$. Therefore, $\mathbf{Pr}(E_j \mid F_j) = \mathbf{Pr}(\hat{N}_{t_j+\tau_{\mathrm{det}}} = N_{t_j+\tau_{\mathrm{det}}} \mid F_j) \geq 1 - \varepsilon$, which completes the proof. $\square$

References

[1] Emmanuel J. Candès, *The restricted isometry property and its implications for compressed sensing*, Comptes Rendus Mathematique, Volume 346, Issues 9–10, May 2008, Pages 589-592, ISSN 1631-073X, http://dx.doi.org/10.1016/j.crma.2008.03.014.