The Mean Square Error in Kalman Filtering Sensor Selection is Approximately Supermodular

Luiz F. O. Chamon,
George J. Pappas and Alejandro Ribeiro

CDC 2017
December 12^{th}, 2017

Why does greedy sensor selection for Kalman filtering works when it shouldn't?

Problem (KFSS)

Select up to s system outputs to estimate its internal states.

$$\begin{aligned} &\underset{\mathcal{S} \subseteq \mathcal{O}}{\text{minimize}} && \text{MSE}(\mathcal{S}) \\ &\text{subject to} && |\mathcal{S}| \leq s \end{aligned}$$

- Why the MSE? KF
- NP-hard [Natarajan'95, Zhang'17, Ye'17]

Greedy KFSS

Definition

Select sensors/outputs one at a time by choosing the one that most improves estimation at each step.

function $\textsc{Greedy}(q)$

$\mathcal{G}_0 = \{\}$

for $j = 1, \dots, q$

$$u = \underset{v \in \mathcal{O} \setminus \mathcal{G}_{j-1}}{\operatorname{argmin}} \ \mathsf{MSE}\left(\mathcal{G}_{j-1} \cup \{v\}\right)$$

$$\mathcal{G}_j = \mathcal{G}_{j-1} \cup \{u\}$$

end

end

Definition

Select sensors/outputs one at a time by choosing the one that most improves estimation at each step.

function $\textsc{Greedy}(q)$

  $\mathcal{G}_0 = \{\}$

  for $j = 1, \ldots, q$

    $u = \underset{v \in \mathcal{O} \setminus \mathcal{G}_{j-1}}{\operatorname{argmin}} \ \mathsf{MSE}\left(\mathcal{G}_{j-1} \cup \{v\}\right)$

    $\mathcal{G}_j = \mathcal{G}_{j-1} \cup \{u\}$

  end

end

- Low complexity
- Sequential
- Near-optimal for supermodular objectives

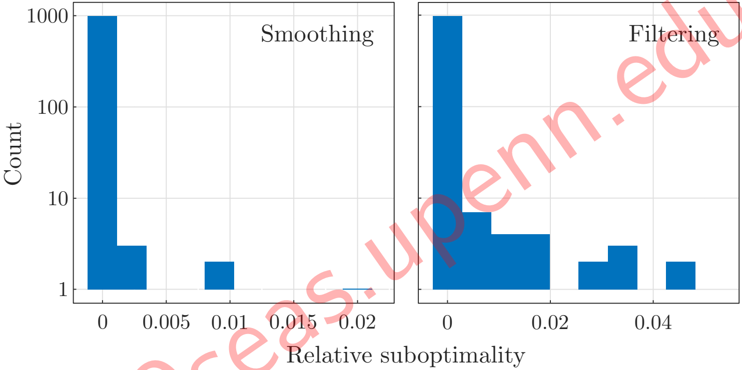
Problem (KFSS)

Select up to s system outputs to estimate its internal states.

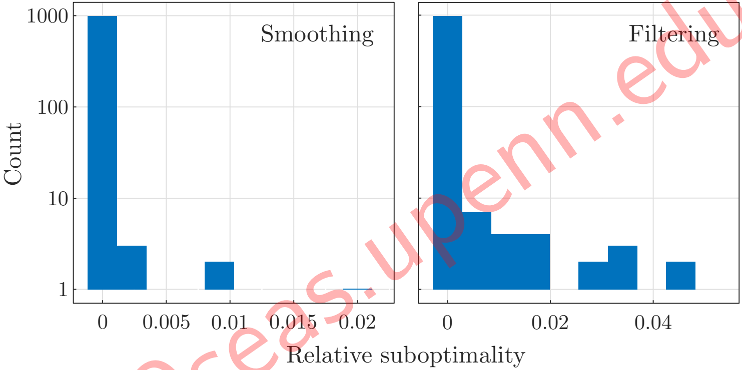
$$\begin{aligned} \underset{\mathcal{S}\subseteq\mathcal{O}}{\text{minimize}} \quad & \text{MSE}(\mathcal{S}) \\ \text{subject to} \quad & |\mathcal{S}| \leq s \end{aligned}$$

- Why the MSE? KF
- NP-hard [Natarajan'95, Zhang'17, Ye'17]
- Estimation MSE is not supermodular [Tzoumas'16, Olshevsky'16, Singh'17, Zhang'17]
  - Use a supermodular surrogate (e.g., $\log\det$) [Joshi'09, Shamaiah'10, Tzoumas'16]

Greedy KFSS



$$\frac{\text{MSE}(\mathcal{G}) - \text{MSE}(\mathcal{S}^\star)}{\text{MSE}(\emptyset) - \text{MSE}(\mathcal{S}^\star)}$$



$$\frac{\mathrm{MSE}(\mathcal{G}) - \mathrm{MSE}(\mathcal{S}^\star)}{\mathrm{MSE}(\emptyset) - \mathrm{MSE}(\mathcal{S}^\star)}$$

Greedy KFSS is near-optimal

Kalman filtering sensor selection

(Approximate) supermodularity

Near-optimality of greedy KFSS

Kalman filtering

$$\boldsymbol{x}_{k+1} = \boldsymbol{F}\boldsymbol{x}_k + \boldsymbol{w}_k$$
$$\boldsymbol{y}_k = \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{v}_k$$
$$\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \sigma_w^2 \boldsymbol{I}) \qquad \boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{0}, \sigma_v^2 \boldsymbol{I}) \qquad \boldsymbol{x}_0 \sim \mathcal{N}(\bar{\boldsymbol{x}}_0, \boldsymbol{\Pi}_0)$$

Problem (Filtering)

Estimate $\boldsymbol{x}_k$ based on outputs up to time k, i.e.,

$$\hat{\boldsymbol{x}}_k = \mathbb{E}\left[\boldsymbol{x}_k \mid \{\boldsymbol{y}_j\}_{j \leq k}\right]$$

Solution (Kalman filter)

$$\hat{\boldsymbol{x}}_k = \boldsymbol{F}\hat{\boldsymbol{x}}_{k-1} + \boldsymbol{K}_k \left[\boldsymbol{y}_k - \boldsymbol{H}\boldsymbol{F}\hat{\boldsymbol{x}}_{k-1}\right]$$

Kalman filtering

$$\boldsymbol{x}_{k+1} = \boldsymbol{F}\boldsymbol{x}_k + \boldsymbol{w}_k$$
$$\boldsymbol{y}_k = \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{v}_k$$
$$\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \sigma_w^2 \boldsymbol{I}) \qquad \boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{0}, \sigma_v^2 \boldsymbol{I}) \qquad \boldsymbol{x}_0 \sim \mathcal{N}(\bar{\boldsymbol{x}}_0, \boldsymbol{\Pi}_0)$$

Problem (Filtering)

Estimate $\boldsymbol{x}_k$ based on outputs in $\mathcal{S} \subseteq \mathcal{O}$ up to time k, i.e.,

$$\hat{\boldsymbol{x}}_k = \mathbb{E}\left[\boldsymbol{x}_k \mid \{(\boldsymbol{y}_j)_{\mathcal{S}}\}_{j \leq k}\right]$$

Solution (Kalman filter)

$$\hat{\boldsymbol{x}}_k(\mathcal{S}) = \boldsymbol{F}\hat{\boldsymbol{x}}_{k-1}(\mathcal{S}) + \boldsymbol{K}_k\left[(\boldsymbol{y}_k)_{\mathcal{S}} - \boldsymbol{H}_{\mathcal{S}}\boldsymbol{F}\hat{\boldsymbol{x}}_{k-1}(\mathcal{S})\right]$$

Problem (KF sensor selection)

Find $\mathcal{S} \subseteq \mathcal{O}$, $|\mathcal{S}| \leq s$, that minimizes the estimation MSE

$$\underset{|\mathcal{S}| \leq s}{\text{minimize}} \quad \sum_{j=0}^{m-1} \theta_j \text{MSE}_{\ell+j}(\mathcal{S})$$

- Myopic sensor selection: $m = 1$
- Final estimation MSE: $\theta_j = 0$ for $j < m - 1$ and $\theta_{m-1} = 1$
- Exponentially weighted error: $\theta_j = \rho^{m-1-j}$, $\rho < 1$

Problem (KF sensor selection)

Find $\mathcal{S} \subseteq \mathcal{O}$, $|\mathcal{S}| \leq s$, that minimizes the estimation MSE

$$\underset{|\mathcal{S}| \leq s}{\text{minimize}} \quad \sum_{j=0}^{m-1} \theta_j \operatorname{Tr}\left[\boldsymbol{P}_{\ell+j}(\mathcal{S})\right]$$

where

$$\boldsymbol{P}_k(\mathcal{S}) = \Bigg(\underbrace{\boldsymbol{F}\boldsymbol{P}_{k-1}(\mathcal{S})\boldsymbol{F}^T + \sigma_w^2 \boldsymbol{I}}_{\boldsymbol{P}_{k|k-1}} + \sigma_v^{-2} \sum_{i \in \mathcal{S}} \underbrace{\boldsymbol{h}_i \boldsymbol{h}_i^T}_{i\text{-th sensor}} \Bigg)^{-1}$$

Kalman filtering sensor selection

(Approximate) supermodularity

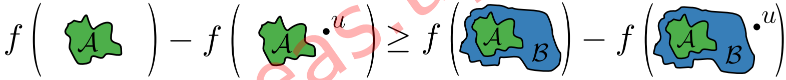
Near-optimality of greedy KFSS

Supermodularity

Definition (Supermodularity)

For $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O}$ and $u \in \mathcal{O} \setminus \mathcal{B}$

$$f\left(\mathcal{A}\right) - f\left(\mathcal{A} \cup \{u\}\right) \geq f\left(\mathcal{B}\right) - f\left(\mathcal{B} \cup \{u\}\right)$$



"diminishing returns"

Greedy supermodular minimization

Theorem ([NWF'78])

Let $\mathcal{S}^\star$ be the optimal solution of the problem

$$\underset{|\mathcal{S}| \leq s}{\text{minimize}} \quad f(\mathcal{S})$$

and $\mathcal{G}$ be its greedy solution. If f is (i) monotone decreasing and (ii) supermodular, then

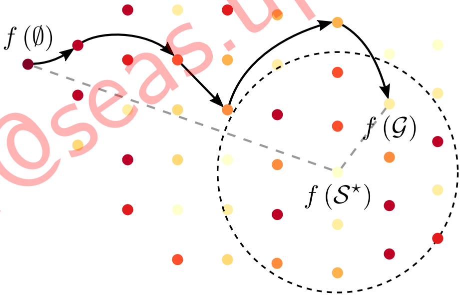
$$\frac{f(\mathcal{G}) - f(\mathcal{S}^\star)}{f(\emptyset) - f(\mathcal{S}^\star)} \leq e^{-1} \approx 0.37.$$

Greedy supermodular minimization

Theorem ([NWF'78])

If f is (i) monotone decreasing and (ii) supermodular, then

$$\frac{f(\mathcal{G}) - f(\mathcal{S}^\star)}{f(\emptyset) - f(\mathcal{S}^\star)} \leq e^{-1} \approx 0.37.$$



Definition (Supermodularity)

For $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O}$ and $u \in \mathcal{O} \setminus \mathcal{B}$

$$f(\mathcal{A} \cup \{u\}) - f(\mathcal{A}) \leq f(\mathcal{B} \cup \{u\}) - f(\mathcal{B})$$

Definition (α-supermodularity)

For $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O}$, $u \in \mathcal{O} \setminus \mathcal{B}$, and $\alpha \geq 0$

$$f(\mathcal{A} \cup \{u\}) - f(\mathcal{A}) \leq \alpha \left[f(\mathcal{B} \cup \{u\}) - f(\mathcal{B}) \right]$$

- If $\alpha \geq 1$: f is supermodular
- If $\alpha < 1$: f is *approximately* supermodular

Greedy α-supermodular minimization

Theorem ([Chamon-Ribeiro'16])

Let $\mathcal{S}^\star$ be the solution of the problem

$$\underset{|\mathcal{S}| \leq s}{\text{minimize}} \quad f(\mathcal{S})$$

and $\mathcal{G}_q$ be the q-th iteration of a greedy solution. If f is (i) monotone decreasing and (ii) α-supermodular, then

$$\frac{f(\mathcal{G}_q) - f(\mathcal{S}^\star)}{f(\emptyset) - f(\mathcal{S}^\star)} \leq e^{-\alpha q/s}.$$

Greedy α-supermodular minimization

Theorem ([Chamon-Ribeiro'16])

If f is (i) monotone decreasing and (ii) α-supermodular, then

$$\frac{f(\mathcal{G}_q) - f(\mathcal{S}^\star)}{f(\emptyset) - f(\mathcal{S}^\star)} \leq e^{-\alpha q/s}.$$

- For $q = s$ and $\alpha = 1$, we recover the classical e^{-1} result
- If $\alpha < 1$, then e^{-1} is recovered for $q = \alpha^{-1}s$

α for KFSS

- What is α for KFSS? Combinatorial problem

$$\alpha = \min_{\substack{\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O} \\ u \in \mathcal{O} \setminus \mathcal{B}}} \frac{\mathsf{MSE}\left(\mathcal{A}\right) - \mathsf{MSE}\left(\mathcal{A} \cup \{u\}\right)}{\mathsf{MSE}\left(\mathcal{B}\right) - \mathsf{MSE}\left(\mathcal{B} \cup \{u\}\right)}$$

α for KFSS

Theorem ([Chamon-Pappas-Ribeiro'17])

The objective of KFSS is α-supermodular with

$$\alpha \geq \min_{\ell \leq k \leq \ell+m-1} \frac{\lambda_{min}\left[\boldsymbol{P}_k(\mathcal{O})\right]}{\lambda_{max}\left[\boldsymbol{P}_{k|k-1}\right]}$$

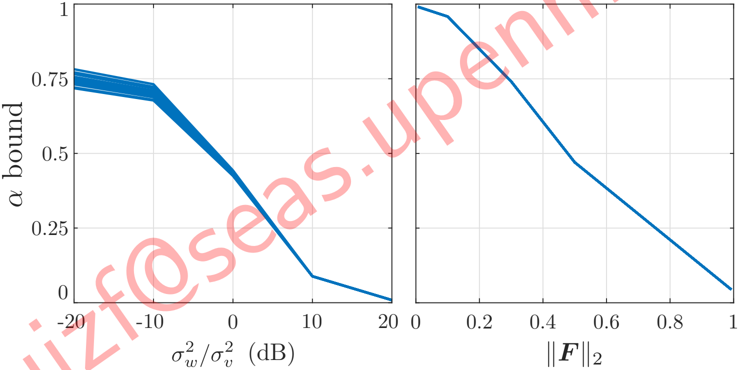
$$\boldsymbol{P}_k(\mathcal{O}) = \left(\boldsymbol{P}_{k|k-1} + \sigma_v^{-2}\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}$$

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{F}\boldsymbol{P}_{k-1}\boldsymbol{F}^T + \sigma_w^2\boldsymbol{I}$$

- $\sigma_v^2 \gg \sigma_w^2$ and small $\kappa(\boldsymbol{F})$ $\Rightarrow \alpha \approx 1$

α for KFSS

- $n = 100$ states and $\boldsymbol{H} = \boldsymbol{I}$



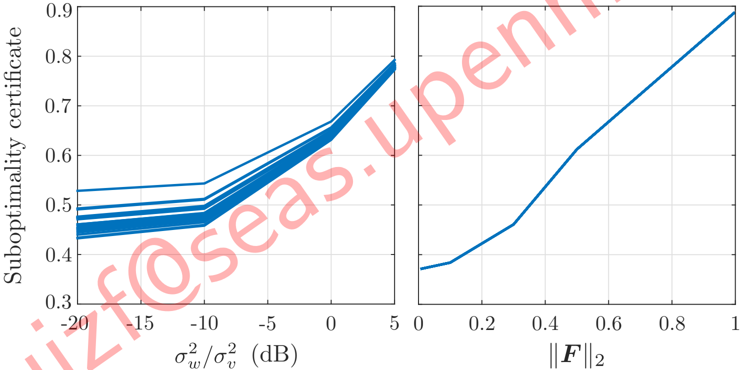
Kalman filtering sensor selection

(Approximate) supermodularity

Near-optimality of greedy KFSS

α for KFSS

- $n = 100$ states and $\boldsymbol{H} = \boldsymbol{I}$



Why does greedy KFSS works so well?

- The MSE in KFSS is not supermodular, but almost
- Greedy KFSS is efficient and has a guaranteed near-optimal performance

The Mean Square Error in Kalman Filtering Sensor Selection is Approximately Supermodular

Luiz F. O. Chamon,
George J. Pappas and Alejandro Ribeiro

More details: http://www.seas.upenn.edu/∼luizf