Center for Science of Information
NSF Science and Technology Center

Minimax Filtering via Relations between Information and Estimation

Albert No
STANFORD University

Introduction

Recent result of information and estimation imply a correspondence between estimation and channel capacity. We apply these results to (causal) filtering of an AWGN-corrupted signal in continuous time.In this poster we focus on an example where the signal is known to be a linear combination of given orthonormal signal set with a power constraint, and we further know that some fraction of coefficients should be zero. In this setting, the corresponding channel capacity problem is that for Gaussian channels with a duty cycle and power constraints, as recently considered in [2].

Problem Setting

- Orthonormal signals set : $\{\phi_i(t), 0 \le t \le T\}_{i=1}^n$
- $X_t = \sum_{i=1}^n c_i \phi_i(t)$
- $C \sim P$ where $P \in \mathcal{P} = \{P : \mathbb{E}_P \|C\|_2^2 \le nA \text{ and } \mathbb{E}_P \|C\|_0 \le k\}$
- $dY_t = X_t dt + dW_t$

Define

- $\mathrm{cmse}_{P,Q} = \mathbb{E}_P \left[\int_0^T (X_t - \mathbb{E}_Q[X_t|Y^t])^2 dt \right]$
- $\mathrm{minimax}(\mathcal{P}) = \min_{\hat{X}_t(\cdot),\, t\in[0,T]} \max_{P\in\mathcal{P}} \left\{ \mathbb{E}_P \left[\int_0^T (X_t - \hat{X}_t(Y^t))^2 dt \right] - \mathrm{cmse}_{P,P} \right\}$

Question

Characterize $\mathrm{minimax}(\mathcal{P})$ and the filter that achieves it.

Equivalent Problem

- $\left\{ \int_0^t \phi_i(s) dY_s \right\}_{i=1}^n$ is a sufficient statistic for estimating X_t.
- Define

$$\tilde{Y}_i(t) = \int_0^t \phi_i(s) dY_s$$
$$\tilde{W}_i(t) = \int_0^t \phi_i(s) dW_s$$
$$\tilde{X}_i(t) = \int_0^t \phi_i(s) X_s ds$$
$$(\Gamma(t))_{i,j} = \int_0^t \phi_i(s)\phi_j(s) ds$$

- Causal estimation is equivalent to following vector estimation problem,

$$\tilde{Y}(t) = \tilde{X}(t) + \tilde{W}(t) = \Gamma(t)A + \tilde{W}(t)$$

where $\tilde{W}(t) \sim \mathcal{N}(0, \Gamma(t))$
- Note that $\Gamma(t)$ does not have to be a full rank matrix. Using eigenvalue decomposition,

$$\Gamma(t) = V(t)\Lambda(t)V(t)^T$$

- Only using nonzero eigenvalues, we can get equivalent formula

$$\Lambda_{\text{eff}}(t)^{-1/2} V_{\text{eff}}(t)^T \tilde{Y}(t) = \Lambda_{\text{eff}}(t)^{1/2} V_{\text{eff}}(t)^T A + \Lambda_{\text{eff}}(t)^{-1/2} V_{\text{eff}}(t)^T \tilde{W}(t)$$

Note $\Lambda_{\text{eff}}(t)^{-1/2} V_{\text{eff}}(t)^T \tilde{W}(t) \sim \mathcal{N}(0, I_{n-m})$

Theorem : Restrict to Bayesian Estimator

Suppose the signal X^T is governed by $P \in \mathcal{P}$. Let $\mathcal{Q}$ denote convex hull of $\mathcal{P}$. Then, for a general loss function,

$$\mathrm{minimax}(\mathcal{P}) = \min_{Q\in\mathcal{Q}} \max_{P\in\mathcal{P}} \{\mathrm{cmle}_{P,Q} - \mathrm{cmle}_{P,P}\}$$

$\mathrm{minimax}(\mathcal{P})$

- Using the recent result of information and estimation[1],

$$\begin{aligned}
\mathrm{minimax}(\mathcal{P}) &= \min_{Q\in\mathcal{Q}} \max_{P\in\mathcal{P}} \mathrm{cmse}_{P,Q} - \mathrm{cmse}_{P,P} \\
&= \min_{Q\in\mathcal{Q}} \max_{P\in\mathcal{P}} D(P_{Y^T} \| Q_{Y^T}) \\
&= \max_{w\in\mathcal{W}} I(\mathcal{P}; Y^T) \\
&= \max_{P\in\mathcal{P}} I(X^T; Y^T)
\end{aligned}$$

- Minimum achieving distribution Q^* is equal to capacity achieving distribution P^*.
- In our example,

$$\mathrm{minimax}(\mathcal{P}) = \max_{P\in\mathcal{P}} I(C; B)$$

where $B = (b_1, \cdots, b_n)^T$ and $b_i = \int_0^T \phi_i(t) dY_t$.
- This problem coincides with the capacity of Gaussian Channels with duty cycle and power constraints.
- Capacity achieving distribution p_d is i.i.d. and discrete.

$$\mathrm{minimax}(\mathcal{P}) = nI(X; Y)$$

where $X \sim p_d$ and $Y = X + N$ is noise corrupted version of X by independent standard Gaussian noise N.

Optimal Causal Estimator

Minimax estimator is a Bayesian estimator assuming prior distribution on C is i.i.d. p_d. Denote this distribution by Q^*, then the optimal causal minimax estiamtor is

$$\hat{X}_t = \mathbb{E}_{Q^*}[X_t | Y^t]$$

We can compute $\hat{X}_t = \mathbb{E}_{Q^*}[X_t|Y^t] = \mathbb{E}_{Q^*}[X_t|\tilde{Y}(t)]$.

Estimators for Comparison

- Maximum likelihood estimator (with/without thresholding)

$$\hat{C} = \left(\Lambda_{\text{eff}}(t)^{1/2} V_{\text{eff}}(t)^T \right)^{\dagger} \Lambda_{\text{eff}}(t)^{-1/2} V_{\text{eff}}(t)^T \tilde{Y}(t) \quad (1)$$

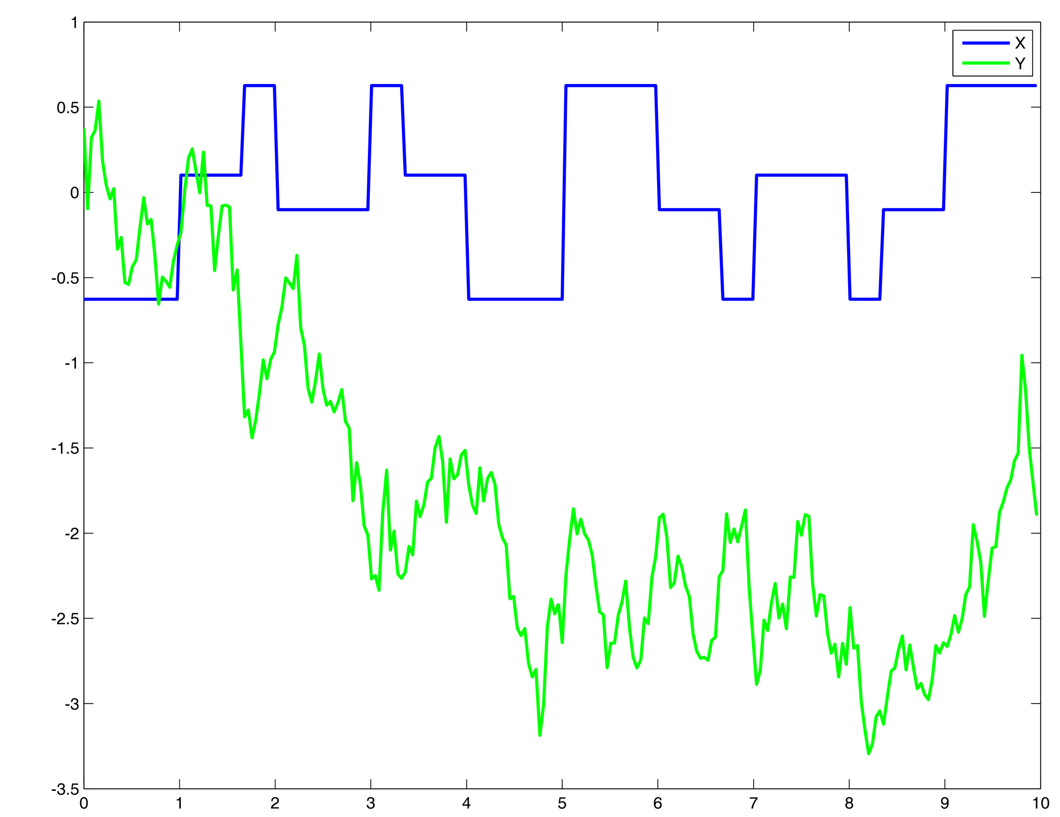
- Minimax estimator that only knows the power constraints

$$\mathbb{E}[C | \Lambda_{\text{eff}}(t)^{-1/2} V_{\text{eff}}(t)^T \tilde{Y}(t)] \quad (2)$$
$$= PV_{\text{eff}}(t) \left(P\Lambda_{\text{eff}}(t) + I_{n-m} \right)^{-1} V_{\text{eff}}(t)^T \tilde{Y}(t) \quad (3)$$

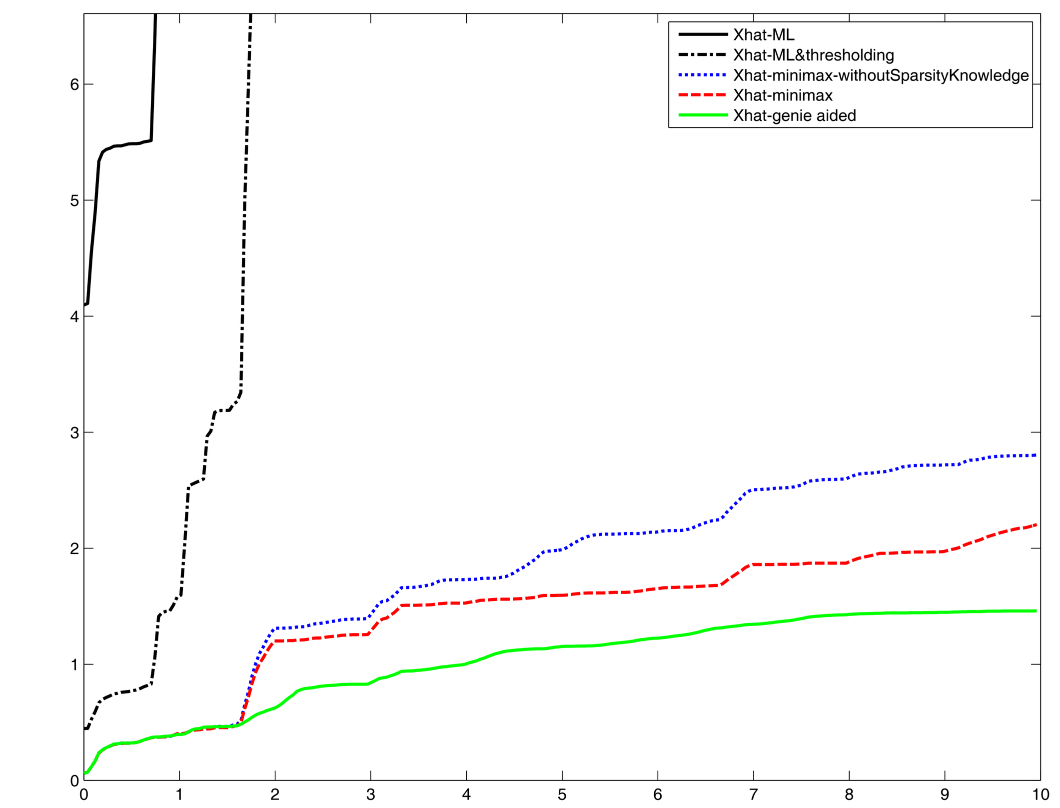
- Genie aided estimator that also knows which coefficients are nonzero

$$\mathbb{E}[C_{\text{nonzero}} | \Lambda_{\text{eff}}(t)^{-1/2} V_{\text{eff}}(t)^T \tilde{Y}(t)] \quad (4)$$
$$= \frac{nP}{k} U_{\text{eff}}^T (U_{\text{eff}} U_{\text{eff}}^T + I_{n-m})^{-1} \Lambda_{\text{eff}}(t)^{-1/2} V_{\text{eff}}(t)^T \tilde{Y}(t) \quad (5)$$

Simulation Results



(a)



(b)

Figure: Simulation Results(Low SNR)

Acknowledgments

The results presented are based on joint work with Tsachy Weissman.

References

- T. Weissman, “The Relationship Between Causal and Noncausal Mismatched Estimation in Continuous-Time AWGN Channels”, IEEE Transactions on Information theory, vol. 56, no. 9, Sep 2010
- Lei Zhang and Dongning Guo, “Capacity of Gaussian Channels with Duty Cycle and Power Constraints”, IEEE Int. Symposium on Information Theory 2011, July, 2011