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: \( \{ \phi_i(t), 0 \leq t \leq T \}_{i=1}^n \)
- \( X_t = \sum_{i=1}^n c_i \phi_i(t) \)
- \( C \sim P \) where \( P \in \mathcal{P} = \{ P : \mathbb{E}[P||C||^2 \leq nA, \mathbb{E}[P||C||_p \leq k] \} \)
- \( dY_t = X_t dt + dW_t \)

Define

\[ \text{cme}_{P,Q} = \mathbb{E}_P \left[ \int_0^T (X_t - \mathbb{E}_{Q}[X_t|Y_t])^2 dt \right] \]

\[ \min_{(P)_{t \in [0,T]}} \max_{\text{cme}_{P,Q}} \left\{ \mathbb{E}_P \left[ \int_0^T (X_t - Xhat_t)^2 dt \right] - \text{cme}_{P,Q} \right\} \]

Question

Characterize \( \min(P) \) and the filter that achieves it.

Equivalent Problem

\[ \left\{ \int_0^T \phi(s) dY_s \right\}_{i=1}^n \] is a sufficient statistic for estimating \( X_t \).

Define

\[ Y_t = \int_0^T \phi(s) dY_s \]
\[ \tilde{W}_t = \int_0^T \phi(s) dW_s \]
\[ \tilde{X}_t = \int_0^T \phi(s) X_s ds \]
\[ (\Gamma(t))_{ij} = \int_0^T \phi(s) \phi(s) ds \]

- Causal estimation is equivalent to following vector estimation problem, \( \tilde{Y}(t) = X_t + \tilde{W}(t) = \Gamma(t)A + \tilde{W}(t) \)

where \( \tilde{W}(t) \sim N(0, \Gamma(t)) \).

- 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

\[ Xeff(t)^{-1/2} Veff(t)^T \tilde{W}(t) = A_{eff}(t)^{-1/2} A_{eff}(t)^T \Lambda(t) + A_{eff}(t)^{-1/2} Veff(t)^T \tilde{W}(t) \]

Note \( A_{eff}(t)^{-1/2} Veff(t)^T \tilde{W}(t) \sim N(0, I_n - m) \)

\[ (\mathcal{C} - \mathbb{E}[\mathcal{C}]) \]

\[ \mathbb{E}[X_{eff}(t)^{-1/2} V_{eff}(t)^T \tilde{Y}(t)] \]

\[ = P_{eff}(t) \left( P_{eff}(t) + I_n - m \right)^{-1} V_{eff}(t)^T \tilde{Y}(t) \]

\[ \mathbb{E}[X_{gen}(t)^{-1/2} V_{eff}(t)^T \tilde{Y}(t)] \]

\[ = nP_{eff} U_{eff} U_{eff}^T + I_n - m \]

\[ \min(P) \] is equal to capacity achieving distribution \( \mathcal{P}^* \).

- This problem coincides with the capacity of Gaussian Channels with duty cycle and power constraints.

- Capacity achieving distribution \( P_{eff} \) is i.i.d. and discrete.

\[ \min_{(P)_{t \in [0,T]}} \max_{\text{cme}_{P,Q}} \left\{ \mathbb{E}_P \left[ \int_0^T (X_t - Xhat_t)^2 dt \right] - \text{cme}_{P,Q} \right\} \]

Optimal Causal Estimator

Minimax estimator is a Bayesian estimator assuming prior distribution on \( C \) is i.i.d. \( P_{eff} \). Denote this distribution by \( \mathcal{Q} \), then the optimal causal minimax estimator is

\[ \hat{X}_t = \mathbb{E}_{\mathcal{Q}}[X_t|Y_t] \]

We can compute \( \hat{X}_t = \mathbb{E}_{\mathcal{Q}}[X_t|Y_t] = \mathbb{E}_{\mathcal{Q}}[X_t|Y_t] \).

Estimators for Comparison

- Maximum likelihood estimator (with/without thresholding)

\[ \tilde{C} = \left( A_{eff}(t)^{-1/2} V_{eff}(t)^T \tilde{Y}(t) \right) \]

\[ \tilde{Y}(t) = A_{eff}(t)^{-1/2} V_{eff}(t)^T \tilde{Y}(t) \]

- Minimax estimator that only knows the power constraints

\[ \mathbb{E}[C|A_{eff}(t)^{-1/2} V_{eff}(t)^T \tilde{Y}(t)] \]

\[ = P_{eff}(t) \left( P_{eff}(t) + I_n - m \right)^{-1} V_{eff}(t)^T \tilde{Y}(t) \]

- genie aided estimator that also knows which coefficients are nonzero

\[ \mathbb{E}[X_{gen}(t)^{-1/2} V_{eff}(t)^T \tilde{Y}(t)] \]

\[ = nP_{eff} U_{eff} U_{eff}^T + I_n - m \]

\[ \min(P) \]

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

References

[1] 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

[2] Le Zhang and Dongning Guo, “Capacity of Gaussian Channels with Duty Cycle and Power Constraints”, IEEE Int. Symposium on Information Theory 2011, July 2011