Fundamental limits of remote estimation of autoregressive Markov processes under communication constraints

Jhelum Chakravorty & Aditya Mahajan

Dept. of ECE, McGill University

Contact Information:
Electrical & Computer Engineering
McGill University
3480 University Street
Montreal, Quebec, Canada H3A 0E9
Email: `jhelum.chakravorty@mail.mcgill.ca`

Motivation

Some applications: Smart grid, environmental monitoring, sensor networks
Salient features: Sensing is cheap, transmission is expensive (e.g. battery-powered transmitter), size of data-packet not critical

Problem formulation

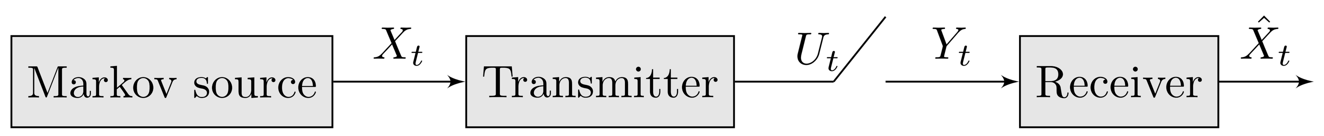

Figure 1: Block diagram of the communication system

- State process: $X_{t+1} = aX_t + W_t$
- Transmitter: $U_t = f_t(X_{0:t}, U_{0:t-1})$;
- *Received symbol*: $Y_t = \begin{cases} X_t, & \text{if } U_t = 1; \\ \mathfrak{E}, & \text{if } U_t = 0, \end{cases}$; Receiver: $\hat{X}_t = g_t(Y_{0:t})$
- **Model A:** $a, X_t, W_t \in \mathbb{Z}$; W_t: with a *unimodal* and *symmetric* distribution p, i.e. for all $e \in \mathbb{Z}_{\geq 0}$, $p_e = p_{-e}$ and $p_e \geq p_{e+1}$; $p_1 > 0$.
- **Model B:** $a, X_t, W_t \in \mathbb{R}$; W_t: zero-mean Gaussian random variable with variance σ^2. The pdf of W_t is denoted by $\phi(\cdot)$.
- *Per-step distortion*: $d(\cdot)$. **Model A**: $d(0) = 0$, for $e \neq 0$, $d(e) \neq 0$; $d(\cdot)$ is *EI* on $\mathbb{Z}_{\geq 0}$, i.e. for all $e \in \mathbb{Z}_{\geq 0}$, $d(e) = d(-e)$ and $d(e) \leq d(e+1)$. **Model B**: $d(e) = e^2$

Performance measures

Discounted case–$\beta \in (0, 1)$:

- Expected distortion:

$$D_{\beta}(f, g) \coloneqq (1-\beta)\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{\infty} \beta^t d(X_t - \hat{X}_t) \;\Big|\; X_0 = 0\Big] \tag{1}$$

- Expected number of transmissions:

$$N_{\beta}(f, g) \coloneqq (1-\beta)\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{\infty} \beta^t U_t \;\Big|\; X_0 = 0\Big] \tag{2}$$

Long-term average–$\beta = 1$: similar to the discounted case. *Take the horizon $T \to \infty$.*

Optimization problems

1. Constrained: Given a discount factor $\beta \in (0, 1]$ and a constant $\alpha \in (0, 1)$, find a transmission and estimation strategy (f^*, g^*) such that
$$D^*_{\beta}(\alpha) \coloneqq D_{\beta}(f^*, g^*) = \inf_{(f,g): N_{\beta}(f,g) \leq \alpha} D_{\beta}(f, g), \tag{3}$$
where the infimum is taken over all history-dependent strategies.
2. Costly (Lagrange relaxation): Given a discount factor $\beta \in (0, 1]$ and a communication cost $\lambda \in \mathbb{R}_{>0}$, find a transmission and estimation strategy (f^*, g^*) such that
$$C^*_{\beta}(\lambda) \coloneqq C_{\beta}(f^*, g^*; \lambda) = \inf_{(f,g)} C_{\beta}(f, g; \lambda), \tag{4}$$
$C_{\beta}(f, g; \lambda) \coloneqq D_{\beta}(f, g) + \lambda N_{\beta}(f, g)$: *total communication cost*. The infimum is taken over all history-dependent strategies.

Main results

Define $Z_t \coloneqq \begin{cases} Y_t, & \text{if } Y_t \neq \mathfrak{E}; \\ aZ_{t-1}, & \text{if } Y_t = \mathfrak{E}, \end{cases}$ and $E_t \coloneqq X_t - Z_{t-1}$.

Structural results

- Optimal estimation strategy: $\hat{X}_t = g^*_t(Z_t) = Z_t$,
- Optimal transmission strategy: $U_t = f_t(E_t) = \begin{cases} 1, & \text{if } |E_t| \geq k; \\ 0, & \text{if } |E_t| < k, \end{cases}$ *time-homogeneous threshold.*

Fix the form of estimator and find *best* transmitter–**centralized** optimization problem. Optimal transmission strategy: *Unique* solution to DP: $\beta \in (0, 1)$,
Model A:
$$V_{\beta}(e; \lambda) = \min_{u \in \{0,1\}} \Big[c(e, u) + \beta \sum_{w \in \mathbb{Z}} p_w V_{\beta}(a(1-u)e + w; \lambda)\Big], \tag{5}$$
where $c(e, u) = (1-\beta)[\lambda u + (1-u)d(e)]$.
Model B: Similar to Model A. *Summation $\to$ integration*
$\beta = 1$: **Vanishing discount approach**–take $\lim \beta \uparrow 1$: Requires: *EI* property of value function, which satisfies certain conditions (SEN).

Performance of threshold-based strategy

Key steps:

1. Define $\mathbf{L}^{(\mathbf{k})}_{\beta}$, $\mathbf{M}^{(\mathbf{k})}_{\beta}$, $\mathbf{D}^{(\mathbf{k})}_{\beta}$, $\mathbf{N}^{(\mathbf{k})}_{\beta}$.
2. Solve for $L^{(k)}_{\beta}(0)$ and $M^{(k)}_{\beta}(0)$.
   Model A: *Closed form matrix-inversion* formula
   Model B: solutions of *balance equations*.
3. *Renewal relationship* to establish expressions for $D^{(k)}_{\beta}(0)$, $N^{(k)}_{\beta}(0)$. $C^{(k)}_{\beta}(0; \lambda) \coloneqq D^{(k)}_{\beta}(0) + \lambda N^{(k)}_{\beta}(0)$

Important features:

- *Monotonicity (and differentiability for Model B)* of $L^{(k)}_{\beta}$, $M^{(k)}_{\beta}$, $D^{(k)}_{\beta}$, $N^{(k)}_{\beta}$.
- $C^{(k)}_{\beta}(0; \lambda)$ is *sub-modular* in (k, λ).

Summary of results for both models

For costly communication:

- Expression for *critical transmission costs* $\lambda^{(k)}_{\beta}$–an increasing sequence.
- For $\lambda \in (\lambda^{(k)}_{\beta}, \lambda^{(k+1)}_{\beta}]$, $f^{(k+1)}$ is optimal.
- $C^*_{\beta}(\lambda)$ is continuous (and piecewise linear for Model A), increasing in λ.

For constrained communication:

- Find optimal threshold $k^*_{\beta}(\alpha)$ corresponding to constraint $\alpha \in (0, 1)$.
- For Model A, $f^*(e) = \begin{cases} 0, & \text{if } |e| < k^*_{\beta}(\alpha); \\ 0, & \text{w.p. } 1-\theta^*, \text{ if } |e| = k^*_{\beta}(\alpha); \\ 1, & \text{w.p. } \theta^*, \text{ if } |e| = k^*_{\beta}(\alpha); \\ 1, & \text{if } |e| > k^*_{\beta}(\alpha). \end{cases}$. For Model B, $f^* = f^{(k^*_{\beta}(\alpha))}$.
- For Model A, compute optimal randomization $\theta^*_{\beta}(\alpha)$.
- For Model A, optimal performance (**DT function**) is *randomized*:
$$D^*_{\beta}(\alpha) = \theta^*_{\beta}(\alpha) D_{\beta}(f^{(k^*_{\beta}(\alpha))}, g^*) + (1-\theta^*_{\beta}(\alpha)) D_{\beta}(f^{(k^*_{\beta}(\alpha)+1)}, g^*). \tag{6}$$
For Model B, $D^*_{\beta}(\alpha) = D^{(k^*_{\beta}(\alpha))}_{\beta}(0)$.
- D^*_{β} is continuous (and piecewise linear for Model A), decreasing and convex in α.

Numerical examples

- Model A: symmetric birth-death Markov chain
- Model B: Gauss-Markov process with $a = 1$.

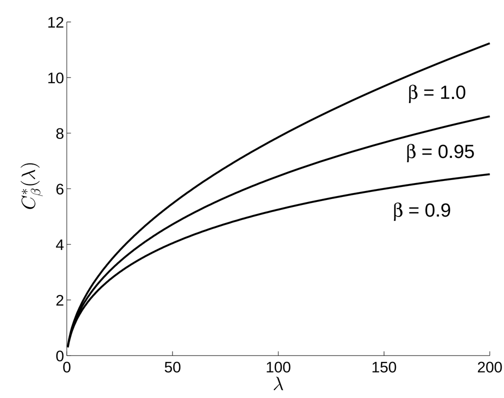

Figure 2: Model B: optimal costly performance for different β

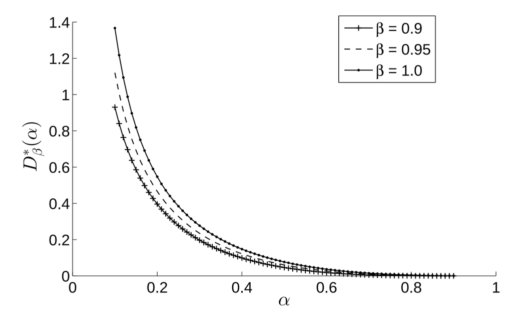

Figure 3: Model B: optimal Constrained performance for different β