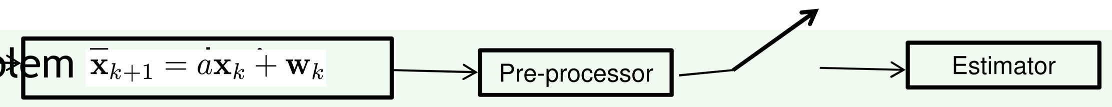
Optimal Distributed State Estimation with Communication Cost: A Majorization Theory Approach

Gabriel M. Lipsa and Nuno C. Martins

Problem



$\bar{\mathbf{x}}_{k+1} = a\mathbf{x}_k + \mathbf{w}_k$

$\mathbf{x}_0 = 0$

$\mathbf{w}_k \sim \mathcal{N}\left(0, \sigma_W^2\right)\left(N_{\sigma_W^2}\right)$

$\mathcal{P}_k : \mathbf{R}^{k+1} \to \mathbf{R}\bigcup\{\emptyset\}$ $\quad \mathbf{v}_k = \mathcal{P}_k\left(\{\mathbf{x}_t\}_{t=0}^k\right)$

$\mathcal{E}_k : (\mathbf{R}\bigcup\{\emptyset\})^{k+1} \to \mathbf{R}$ $\quad \hat{\mathbf{x}}_k = \mathcal{E}_k\left(\{\mathbf{v}_t\}_{t=0}^k\right)$

$$r_k = \begin{cases} 0, & v_k = \emptyset \\ 1, & v_k \neq \emptyset \end{cases}$$

$$\mathcal{J}_\beta\left(\alpha, \sigma_W^2, C, \mathcal{P}, \mathcal{E}\right) = \lim_{T\to\infty} \sum_{k=0}^{T} \beta^k E\left[(\mathbf{x}_k - \hat{\mathbf{x}}_k)^2 + C\mathbf{r}_k\right]$$

$$\mathcal{J}_\beta\left(\alpha, \sigma_W^2, C\right) = \min_{\mathcal{P},\mathcal{E}} \mathcal{J}_\beta\left(\alpha, \sigma_W^2, C, \mathcal{P}, \mathcal{E}\right)$$

Optimal solution:

$$\mathcal{E}_{(\alpha),k}\left(\{v_t\}_{t=0}^k\right) = \hat{x}_k = \begin{cases} \alpha\hat{x}_{k-1}, & v_k = \emptyset \\ v_k, & \text{ow} \end{cases}, \quad k \geq 1$$

$$\mathcal{P}_{(\tau),k}\left(\{x_t\}_{t=0}^k\right) = \begin{cases} \emptyset, & -\tau_k < x_k - \hat{x}_k < \tau_k \\ x_k & \text{ow} \end{cases}, k \geq 1$$

Theorem 1: Let the following parameters be given: the variance of the process noise σ_W^2, the system's dynamic constant α, the communication cost C, and the discount factor β. There exist a positive real constant ν, such that $\mathcal{P}_{(\tau)}$ and $\mathcal{E}_{(\alpha)}$ are an optimal solution to $\min_{\mathcal{P},\mathcal{E}} \mathcal{J}_\beta\left(\alpha, \sigma_W^2, C, \mathcal{P}, \mathcal{E}\right)$, where $\tau = \{\tau_k\}_{k=0}^\infty$ is a sequence of positive real numbers such that $\tau_k = \nu$ for all positive integers k.

$$\mathcal{J}_{T,\beta}\left(\alpha, \sigma_W^2, C, \mathcal{P}, \mathcal{E}\right) = \sum_{k=0}^{T} \beta^k E\left[(\mathbf{x}_k - \hat{\mathbf{x}}_k)^2 + C\mathbf{r}_k\right]$$

$$\mathcal{J}_{T,\beta}\left(\alpha, \sigma_W^2, C\right) = \min_{\mathcal{P},\mathcal{E}} \mathcal{J}_{T,\beta}\left(\alpha, \sigma_W^2, C, \mathcal{P}, \mathcal{E}\right)$$

Theorem 2: Let the following parameters parameters be given: the variance of the process noise σ_W^2, the system's dynamic constant α, the communication cost C, the discount factor β and the time horizon T. There exists a sequence of positive real numbers $\tau = \{\tau_k\}_{k=0}^T$, such that $\mathcal{P}_{(\tau)}$ and $\mathcal{E}_{(\alpha)}$ are an optimal solution to $\min_{\mathcal{P},\mathcal{E}} \mathcal{J}_{T,\beta}\left(\alpha, \sigma_W^2, C, \mathcal{P}, \mathcal{E}\right)$.

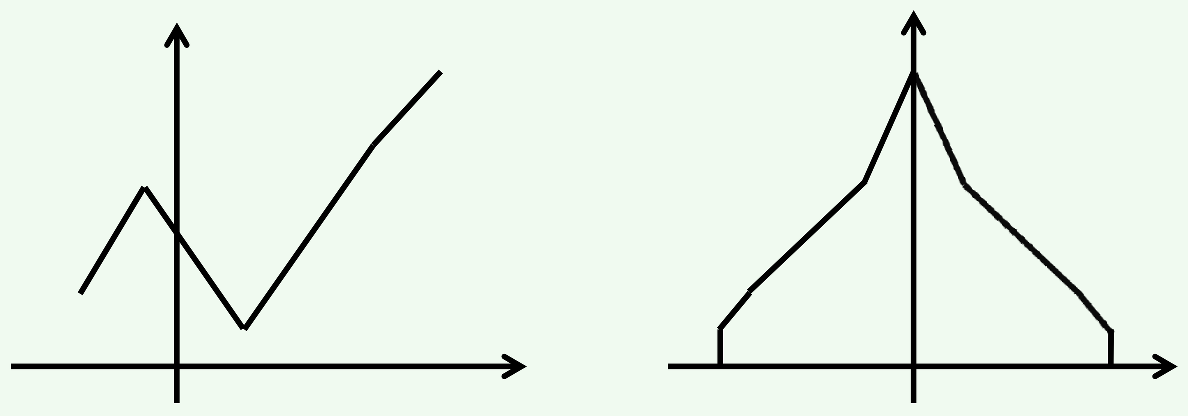
Remark: The optimal cost $J_{T,\beta}\left(\alpha, \sigma_W^2, C\right)$ does not depend on the initial condition $\mathbf{x}_0$.

Remark: For a given pre-processor policy $\mathcal{P}$, the estimtor which minimizes the cost $J_{T,\beta}\left(\alpha, \sigma_W^2, C\right)$ is the conditional expectation of the state $\mathbf{x_k}$ given the entire past, i.e. the estimator which computes $\hat{x}_k = E\left[\mathbf{x}_k | v_t, 1 \leq t \leq k\right]$.

Majorization Theory:

The symmetric nonincreasing rearrangement of the function f is given by:

$$f^\sigma(x) = \int_0^\infty \mathcal{I}_{(z\in\mathbf{R}^n : f(z) > \rho)^\sigma}(x) d\rho$$



$f \qquad\qquad f^\sigma$

$$f \succ g \text{ if: } \int_{||x||\leq\rho} f^\sigma(x)dx \geq \int_{||x||\leq\rho} g^\sigma(x)dx, \text{for all } \rho \geq 0$$

$$f_{\mathbf{K}}(x) = \begin{cases} \dfrac{f(x)}{\int_{\mathbf{K}} f(x)dx}, x \in \mathbf{K} \\ 0, x \notin \mathbf{K} \end{cases}$$

A nonnegative function $f : \mathbf{R}^n \to \mathbf{R}_+$ is symmetric and nonincreasing if it can be expressed via a nonincreasing function $g : \mathbf{R}_+ \to \mathbf{R}_+$ such that $f(x) = g(||x||)$.

Lemma: Let $f, g : \mathbf{R} \to \mathbf{R}$ be two probability distribution functions, such that f is symmetric and nonincreasing and $f \succ g$. Let κ be a real number in the interval $\kappa \in (0,1)$. Let $\mathbf{K} = [-\tau, \tau]$ be the symmetric interval, such that $\int_{-\tau}^{\tau} f(x)dx = 1 - \kappa$. For any function $\lambda : \mathbf{R} \to [0,1]$ satisfying $\int_{\mathbf{R}} g(x)\lambda(x)dx = 1 - \kappa$, the following holds:

$$f_{\mathbf{K}} \succ \frac{g\cdot\lambda}{1-\kappa}$$

Lemma [Hajek]: Let f and g be two probability distribution functions on $\mathbf{R}^n$, with f is symmetric and nonincreasing and $f \succ g$. For a symmetric nonincreasing probability distribution function b the following holds:

$$f * b \succ g * b$$

Lemma: Let f be a neat and even probability distribution function on the real line. Let μ, be a probability distribution function on the real line, such that $\mu \prec f$. If we define $\hat{x}_\mu = \int_{\mathbf{R}} x\mu(x)dx$, then the following holds:

$$\int_{\mathbf{R}} x^2 f(x)dx \leq \int_{\mathbf{R}} (x - \hat{x}_\mu)^2 \mu(x)dx$$

Preliminary notations:

$\overline{\omega}_k = f\left(x_k | \mathbf{v}_t = \emptyset, 1 \leq t \leq k, x_0\right)$

$\gamma_k = P\left(v_t = \emptyset, 1 \leq t \leq k, x_0\right)$

$\gamma_{k|k-1} = P\left(v_k = \emptyset | v_t = \emptyset, 1 \leq t \leq k-1, x_0\right)$

Proof of Theorem 2:

Theorem 2 can be proved using an induction argument. Choose an arbitrary pre-processor policy $\mathcal{P}$, this policy will define $\overline{\omega}_k$, $\gamma_{k|k-1}$ and γ_k for all $k \in \{1, \ldots, T\}$

Let $\{\tau_k\}_{k=1}^T$ be a sequence of positive real numbers. We define the policy $\mathcal{P}^o$ as it follows:

- if $\mathbf{v}_t = \emptyset$ for all $t \in \{1, \ldots, k-1\}$, let $v_k = \emptyset$ if $\mathbf{x}_k \in [-\tau_k, \tau_k]$, otherwise let $\mathbf{v}_k = \mathbf{x}_k$;
- after a perfect sample was sent, adopt the optimal threshold policy for the remaining time.

The policy $\mathcal{P}^o$ will define $\overline{\omega}_k^o$, $\gamma_{k|k-1}^o$ and γ_k^o for all $k \in \{1, \ldots, T\}$. Choose the thresholds τ_k such that $\gamma_{k|k-1} = \gamma_{k|k-1}^o$ for all $k \in \{1, \ldots, T\}$. It can be shown that $\overline{\omega}_k^o \succ \overline{\omega}_k$ for all $k \in \{1, \ldots, T\}$, and moreover $\overline{\omega}_k^o$ is symmetric and nonincreasing.

The cost obtained by adopting the policy $\mathcal{P}$ can be lower bounded as follows:

$$\mathcal{J}_{T,\beta}\left(\alpha, \sigma_W^2, C, \mathcal{P}, \mathcal{E}\left(\mathcal{P}\right)\right) \geq \textstyle\sum_{k=1}^T \Big(E_{\overline{\omega}_k}\left[(\mathbf{x}_k - \hat{x}_k)^2\right]\gamma_k + \left(C + \mathcal{J}_{T-k,\beta}\left(\alpha, \sigma_W^2, C\right)\right)(1 - \gamma_{k|k-1})\gamma_{k-1}\Big)$$

The cost obtained by adopting the policy $\mathcal{P}^o$ is given by:

$$\mathcal{J}_{T,\beta}\left(\alpha, \sigma_W^2, C, \mathcal{P}^o, \mathcal{E}\left(\mathcal{P}^o\right)\right) = \textstyle\sum_{k=1}^T \Big(E_{\overline{\omega}_k^o}\left[(\mathbf{x}_k - \hat{x}_k^o)^2\right]\gamma_k + \left(C + \mathcal{J}_{T-k,\beta}\left(\alpha, \sigma_W^2, C\right)\right)(1 - \gamma_{k|k-1})\gamma_{k-1}\Big)$$

It follows then that:

$$\mathcal{J}_{T,\beta}\left(\alpha, \sigma_W^2, C, \mathcal{P}, \mathcal{E}\left(\mathcal{P}\right)\right) \geq \mathcal{J}_{T,\beta}\left(\alpha, \sigma_W^2, C, \mathcal{P}^o, \mathcal{E}\left(\mathcal{P}^o\right)\right)$$

Proof of Theorem 1:

Let the time horizon T go to infinity and the results follow.

Conclusions:

We have shown that for the given problem, the decision to send or not to send a sample of the underlying stochastic process depends only on the estimation error and that there exists an optimal threshold policy.

References:

Bruce Hajek, Kevin Mitzel and Sichao Yang, *"Paging and Registration in Cellular Networks: Jointly Optimal Policies and an Iterative Algorithm,"* IEEE Transactions on Information Theory, Feb 2008

O. C. Imer and T. Basar. *"Optimal estimation with limited measurements,"* in Proc. IEEE CDC/ECC 2005, Seville, Spain submitted to IEEE Transactions on Automatic Control

Albert W. Marshall, Ingram Olkin, Barry Arnold, *"Inequalities: Theory of Majorization and Its Applications (Springer Series in Statistics),"* Academic Press, New York