

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

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

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

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

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

Optimal solution:

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

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

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 $\min_{\mathcal{P}, \mathcal{E}} \mathcal{J}_\beta(\alpha, \sigma_W^2, C, \mathcal{P}, \mathcal{E})$, where $\tau = \{\tau_k\}_{k=0}^\infty$ is a sequence of positive real numbers such that $\tau_k = \nu$ for all positive integers k .

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

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

Theorem 2: 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 , 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 $\min_{\mathcal{P}, \mathcal{E}} \mathcal{J}_{T, \beta}(\alpha, \sigma_W^2, C, \mathcal{P}, \mathcal{E})$.

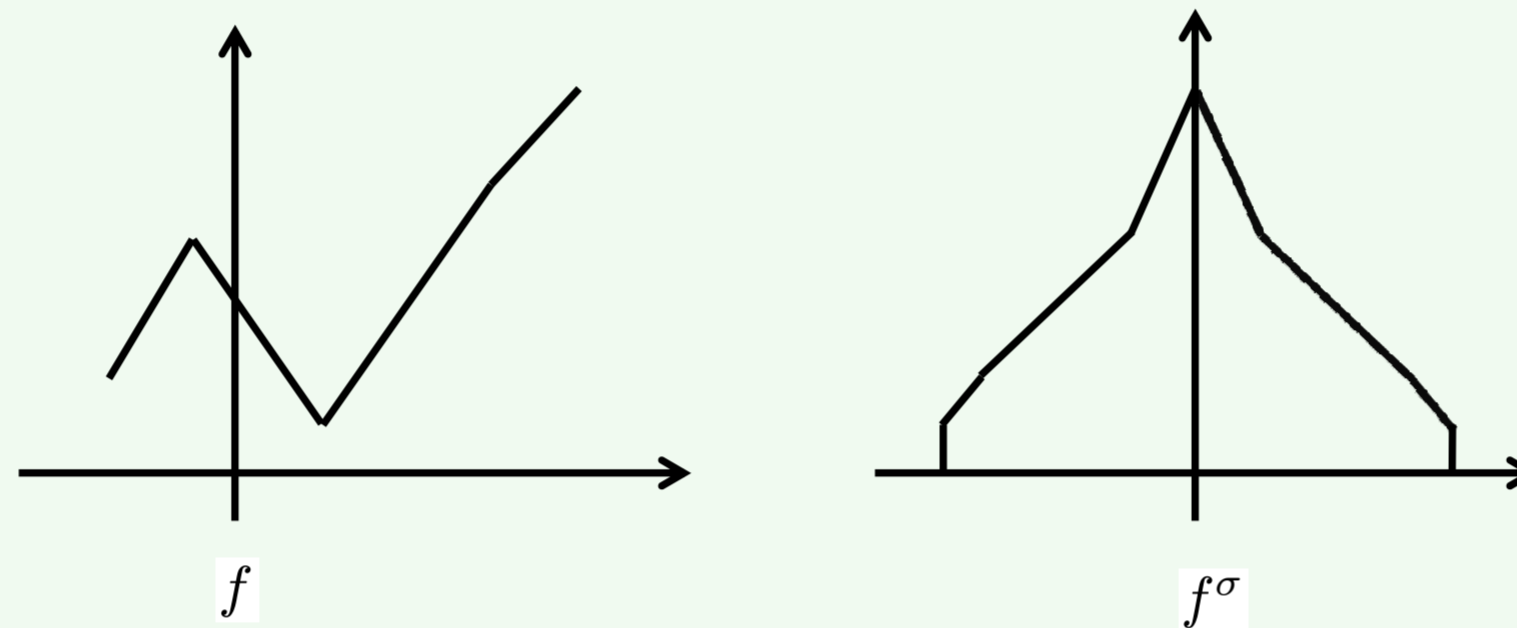
Remark: The optimal cost $\mathcal{J}_{T, \beta}(\alpha, \sigma_W^2, C)$ does not depend on the initial condition \mathbf{x}_0 .

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

Majorization Theory:

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

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



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

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

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

Lemma: Let $f, g : \mathbf{R} \rightarrow \mathbf{R}$ 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 $\lambda : \mathbf{R} \rightarrow [0, 1]$ 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:

$$\bar{\omega}_k = f(x_k | v_t = \emptyset, 1 \leq t \leq k, x_0)$$

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

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

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 $\bar{\omega}_k$, $\gamma_{k|k-1}$ and γ_k for all $k \in \{1, \dots, T\}$

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

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

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

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

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

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

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

It follows then that:

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

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