Distortion-transmission trade-off in real-time transmission of Gauss-Markov sources

Jhelum Chakravorty, Aditya Mahajan

McGill University

IEEE International Symposium on Information Theory, HK, June 14-19, 2015

Motivation

- Sequential transmission of data
- Zero delay in reconstruction

Motivation

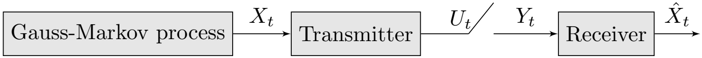
- Sequential transmission of data
- Zero delay in reconstruction

Applications

- Smart grids
- Environmental monitoring
- Sensor networks

- Sensing is cheap
- Transmission is expensive
- Size of data-packet is not critical

The remote-state estimation setup



Source process $X_{t+1} = X_t + W_t$, $W_t \sim \mathcal{N}(0, \sigma^2)$, i.i.d.
Uncontrolled Gauss-Markov process.

Transmitter $U_t = f_t(X_{1:t}, U_{1:t-1})$ and $Y_t = \begin{cases} X_t, & \text{if } U_t = 1; \\ \mathfrak{E}, & \text{if } U_t = 0, \end{cases}$

Receiver $\hat{X}_t = g_t(Y_{1:t})$
Distortion: $(X_t - \hat{X}_t)^2$

Communication strategies
Transmission strategy $f = \{f_t\}_{t=0}^{\infty}$
Estimation strategy $g = \{g_t\}_{t=0}^{\infty}$

The optimization problem

- $D(f,g) := \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}^{(f,g)} \Big[\sum_{t=0}^{T-1} d(X_t - \hat{X}_t) \,\Big|\, X_0 = 0 \Big]$
- $N(f,g) := \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}^{(f,g)} \Big[\sum_{t=0}^{T-1} U_t \,\Big|\, X_0 = 0 \Big]$

The optimization problem

- $D(f, g) := \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{T-1} d(X_t - \hat{X}_t) \Bigm| X_0 = 0\Big]$
- $N(f, g) := \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{T-1} U_t \Bigm| X_0 = 0\Big]$

The Distortion-Transmission function

$$D^*(\alpha) := D(f^*, g^*) := \inf_{(f,g):N(f,g)\le\alpha} D(f, g)$$

Minimize expected distortion such that expected number of transmissions is less than α

Literature overview

Costly communication: analysis of optimal performance

- Estimation with measurement cost: estimator decides whether the sensor should transmit - Athans, 1972; Geromel, 1989; Wu et al, 2008.
- Sensor sleep scheduling: sensor is allowed to sleep for a pre-specified amount of time - Shuman and Liu, 2006; Sarkar and Cruz, 2004, 2005; Federgruen and So, 1991.
- Censoring sensors: sequential hypothesis testing setup; sensor decides whether to transmit or not - Rago et al, 1996; Appadwedula et al, 2008.

Literature overview

Remote state estimation: focus on structure of optimal strategies

- Gauss-Markov source with finite number of transmissions - Imer and Basar, 2005.
- Gauss-Markov source with costly communication (finite horizon) - Lipsa and Martins, 2011; Molin and Hirche, 2012; Xu and Hespanha, 2004.
- Countable Markov source with costly communication (finite horizon) - Nayyar et al, 2013.

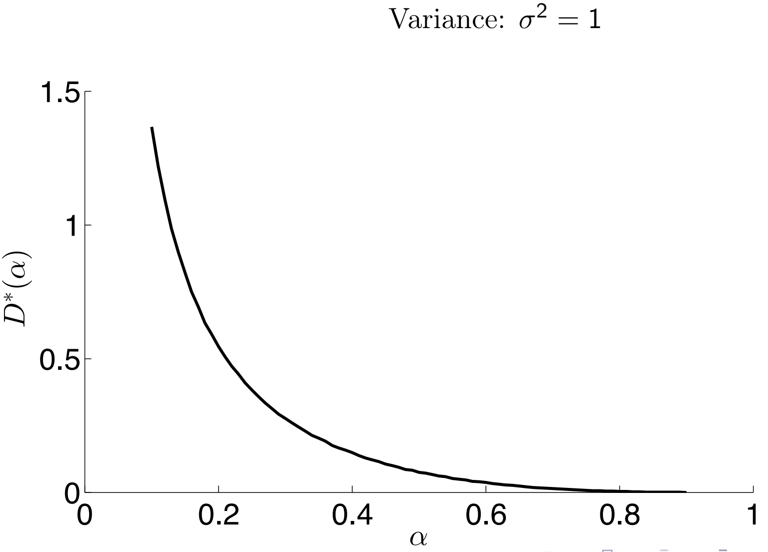
Literature overview

Remote state estimation: focus on structure of optimal strategies

- Gauss-Markov source with finite number of transmissions - Imer and Basar, 2005.
- Gauss-Markov source with costly communication (finite horizon) - Lipsa and Martins, 2011; Molin and Hirche, 2012; Xu and Hespanha, 2004.
- Countable Markov source with costly communication (finite horizon) - Nayyar et al, 2013.

Gauss-Markov source; infinite horizon setup; constrained optimization.

Main result: the Distortion-Transmission function



Main result: the Distortion-Transmission function

How to compute $D^*(\alpha)$ for a given $\alpha \in (0,1)$?

Main result: the Distortion-Transmission function

How to compute $D^*(\alpha)$ for a given $\alpha \in (0,1)$?

- Find $k^*(\alpha) \in \mathbb{R}_{\geq 0}$ such that $M^{(k^*(\alpha))}(0) = 1/\alpha$, where
$$M^{(k)}(e) = 1 + \int_{-k}^{k} \phi(w-e) M^{(k)}(w) dw.$$

- Compute $L^{(k^*(\alpha))}(0)$ where
$$L^{(k)}(e) = e^2 + \int_{-k}^{k} \phi(w-e) L^{(k)}(w) dw.$$

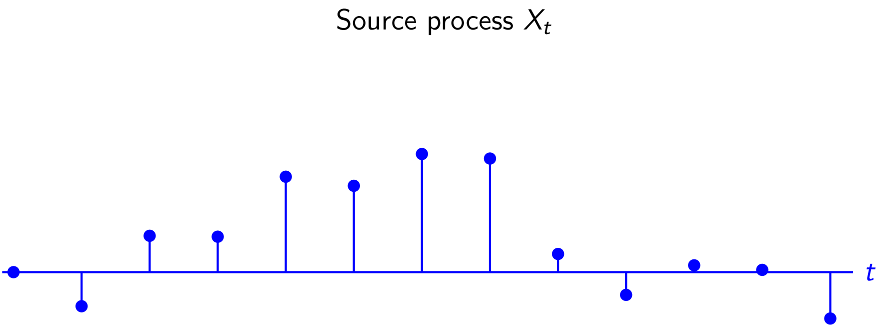
- $D^*(\alpha) = L^{(k^*(\alpha))}(0) / M^{(k^*(\alpha))}(0)$.

- Scaling of distortion-transmission function with variance.
$$D^*_{\sigma}(\alpha) = \sigma^2 D^*_{1}(\alpha).$$

An illustration

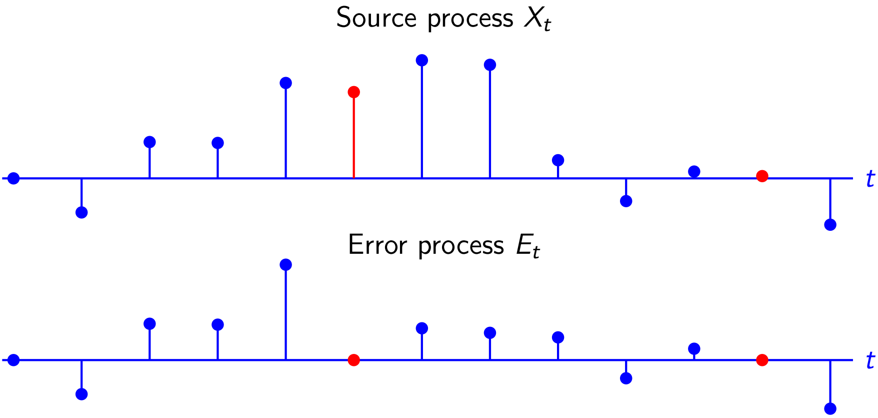
Comparison with periodic strategy

An illustration



An illustration

$\alpha = 1/6$, Periodic strategy

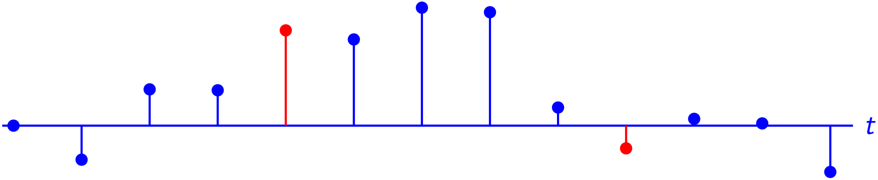


Distortion = 2.083

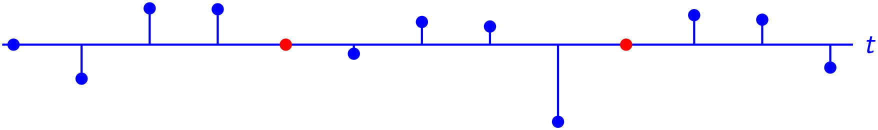
An illustration

$\alpha = 1/6$, Threshold strategy; Threshold=2

Source process X_t

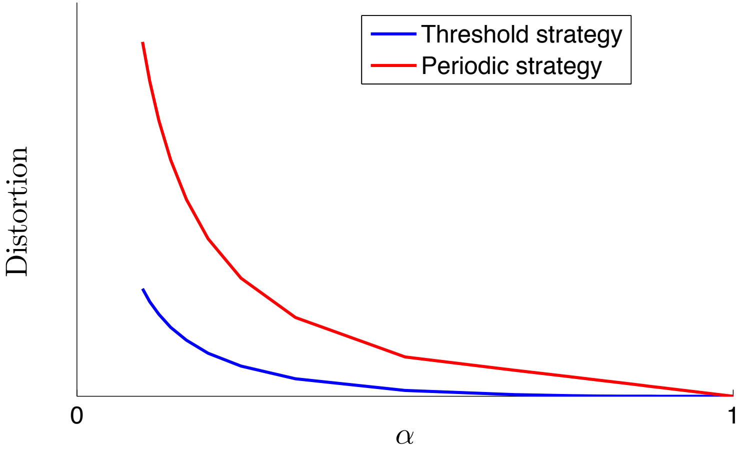


Error process E_t



Distortion = 1.5

An illustration



Proof outline

We don't proceed in the usual way to find the achievable scheme and a converse ! Instead,

Proof outline

We don't proceed in the usual way to find the achievable scheme and a converse ! Instead,

- Identify structure of optimal strategies.
- Find the best strategy with that structure.

Lagrange relaxation

$$C^*(\lambda) := \inf_{(f,g)} C(f, g; \lambda),$$

where $C(f, g; \lambda) = D(f, g) + \lambda N(f, g)$, $\lambda \geq 0$.

Structure of optimal strategies

The structure of optimal transmitter and estimator follows from [Lipsa-Martins 2011] and [Nayyar-Basar-Teneketzis-Veeravalli 2013].

Finite horizon setup; results for Lagrange relaxation

Optimal estimation strategy Let Z_t be the most recently transmitted symbol. $\hat{X}_t = g_t^*(Z_t) = Z_t$; Time homogeneous!

Optimal transmission strategy Let $E_t = X_t - Z_{t-1}$ be the error process and f_t be the threshold based strategy such that

$$f_t(X_t, Y_{0:t-1}) = \begin{cases} 1, & \text{if } |E_t| \geq k_t \\ 0, & \text{if } |E_t| < k_t. \end{cases}$$

Structure of optimal strategies

The structure of optimal transmitter and estimator follows from [Lipsa-Martins 2011] and [Nayyar-Basar-Teneketzis-Veeravalli 2013].

Finite horizon setup; results for Lagrange relaxation

Optimal estimation strategy Let Z_t be the most recently transmitted symbol. $\hat{X}_t = g_t^*(Z_t) = Z_t$; Time homogeneous!

Optimal transmission strategy Let $E_t = X_t - Z_{t-1}$ be the error process and f_t be the threshold based strategy such that

$$f_t(X_t, Y_{0:t-1}) = \begin{cases} 1, & \text{if } |E_t| \geq k_t \\ 0, & \text{if } |E_t| < k_t. \end{cases}$$

We prove that the results generalize to infinite horizon setup; the optimal thresholds are time - homogeneous.

Performance of threshold based strategies

Fix a threshold based startegy $f^{(k)}$. Define

- $D^{(k)}$: the expected distortion.
- $N^{(k)}$: the expected number of transmissions.

Performance of threshold based strategies

Fix a threshold based startegy $f^{(k)}$. Define

- $D^{(k)}$: the expected distortion.
- $N^{(k)}$: the expected number of transmissions.

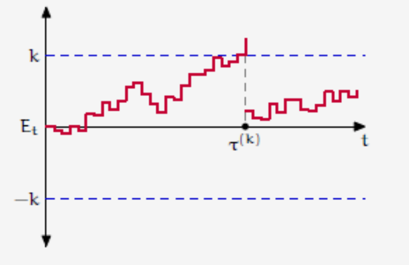
$\{E_t\}_{t=0}^{\infty}$ is regenerative process.

Performance of threshold based strategies

Fix a threshold based startegy $f^{(k)}$. Define

- $D^{(k)}$: the expected distortion.
- $N^{(k)}$: the expected number of transmissions.

$\{E_t\}_{t=0}^{\infty}$ is regenerative process.



$\tau^{(k)}$: stopping time when the Gauss-Markov process starting at state 0 at time $t = 0$ enters the set $\{e \in \mathbb{R} : |e| \geq k\}$

Performance of threshold based strategies

Fix a threshold based startegy $f^{(k)}$. Define

- $D^{(k)}$: the expected distortion.
- $N^{(k)}$: the expected number of transmissions.

$\{E_t\}_{t=0}^{\infty}$ is regenerative process.

- $L^{(k)}(e)$: the expected distortion until the first transmission, starting from state e.
- $M^{(k)}(e)$: the expected time until the first transmission, starting from state e.

Performance of threshold based strategies

Fix a threshold based startegy $f^{(k)}$. Define

- $D^{(k)}$: the expected distortion.
- $N^{(k)}$: the expected number of transmissions.

$\{E_t\}_{t=0}^{\infty}$ is regenerative process.

- $L^{(k)}(e)$: the expected distortion until the first transmission, starting from state e.
- $M^{(k)}(e)$: the expected time until the first transmission, starting from state e.

Renewal relationship

$$D^{(k)} = \frac{L^{(k)}(0)}{M^{(k)}(0)}, \quad N^{(k)} = \frac{1}{M^{(k)}(0)}$$

Performance of threshold based strategies

$$L^{(k)}(e) = e^2 + \int_{-k}^{k} \phi(w-e) L^{(k)}(w) dw;$$
$$M^{(k)}(e) = 1 + \int_{-k}^{k} \phi(w-e) M^{(k)}(w) dw.$$

- Derived using balance equations.
- Solutions of Fredholm Integral Equations of second kind.

Performance of threshold based strategies

$$L^{(k)}(e) = e^2 + \int_{-k}^{k} \phi(w - e) L^{(k)}(w) dw;$$
$$M^{(k)}(e) = 1 + \int_{-k}^{k} \phi(w - e) M^{(k)}(w) dw.$$

- Derived using balance equations.
- Solutions of Fredholm Integral Equations of second kind.

Contraction. Use Banach fixed point theorem to show that

- Fredholm Integral Equations have a solution.
- the solution is unique.

Performance of threshold based strategies

$$L^{(k)}(e) = e^2 + \int_{-k}^{k} \phi(w-e) L^{(k)}(w) dw;$$
$$M^{(k)}(e) = 1 + \int_{-k}^{k} \phi(w-e) M^{(k)}(w) dw.$$

- Derived using balance equations.
- Solutions of Fredholm Integral Equations of second kind.

Computation

- Well-studied numerical methods.
- Examples - use the resolvent kernel of the integral equation - the Liouville-Neumann series; use quadrature method to discretize the integral.

Main theorem

Properties

- $L^{(k)}$, $M^{(k)}$, $D^{(k)}$ and $N^{(k)}$ are continuous, differentiable in k.
- $L^{(k)}$, $M^{(k)}$ and $D^{(k)}$ monotonically increasing in k.
- $N^{(k)}$ is strictly monotonically decreasing in k.

Main theorem

Properties

- $L^{(k)}$, $M^{(k)}$, $D^{(k)}$ and $N^{(k)}$ are continuous, differentiable in k.
- $L^{(k)}$, $M^{(k)}$ and $D^{(k)}$ monotonically increasing in k.
- $N^{(k)}$ is strictly monotonically decreasing in k.

Theorem

- For any $\alpha \in (0,1)$, $\exists k^*(\alpha) : N^{(k^*(\alpha))} = \alpha$.
- If the pair (λ, k), $\lambda, k \in \mathbb{R}_{\geq 0}$, satisfies $\lambda = -\frac{\partial_k D^{(k)}}{\partial_k N^{(k)}}$, then $C^*(\lambda) = C(f^{(k)}, g^*; \lambda)$.
- $D^*(\alpha) = D^{(k^*(\alpha))}$.

Scaling with variance

$$L_{\sigma}^{(k)}(e) = \sigma^2 L_1^{(k/\sigma)}\left(\frac{e}{\sigma}\right), \quad M_{\sigma}^{(k)}(e) = M_1^{(k/\sigma)}\left(\frac{e}{\sigma}\right),$$

Scaling with variance

$$L_{\sigma}^{(k)}(e) = \sigma^2 L_1^{(k/\sigma)}\left(\frac{e}{\sigma}\right), \quad M_{\sigma}^{(k)}(e) = M_1^{(k/\sigma)}\left(\frac{e}{\sigma}\right),$$

Scaling: distortion-transmission function

$$D_{\sigma}^{*}(\alpha) = \sigma^2 D_1^{*}(\alpha).$$

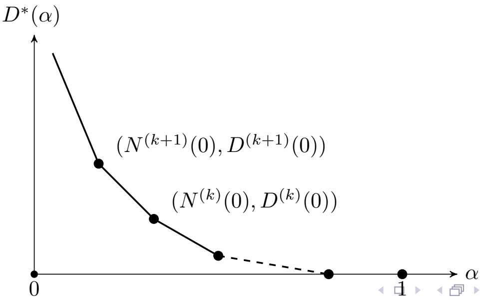
Summary

- Remote state estimation of a Gauss-Markov source under constraints on the number of transmissions.
- Computable expression for distortion-transmission function.
- Simple threshold based strategies are optimal !

Summary

Countable-state Markov chain setup

- Similar results hold - Kalman-like estimator is optimal.
- Randomized threshold based transmission strategy is optimal.
- Distortion-transmission function is piecewise linear, decreasing, convex.



Summary

Countable-state Markov chain setup

- Similar results hold - Kalman-like estimator is optimal.
- Randomized threshold based transmission strategy is optimal.
- Distortion-transmission function is piecewise linear, decreasing, convex.

JC and AM, "Distortion-transmission trade-off in real-time transmission of Markov sources", ITW 2015.

Future directions

- The results are derived under an idealized system model.
- When the transmitter does transmit, it sends the complete state of the source.
- The channel is noiseless and does not introduce any delay.

Future directions

- The results are derived under an idealized system model.
- When the transmitter does transmit, it sends the complete state of the source.
- The channel is noiseless and does not introduce any delay.

Future directions

- Effects of quantization, channel noise and delay.

Future directions

- The results are derived under an idealized system model.
- When the transmitter does transmit, it sends the complete state of the source.
- The channel is noiseless and does not introduce any delay.

Future directions

- Effects of quantization, channel noise and delay.

http://arxiv.org/abs/1505.04829

Some parameters

Let $\tau^{(k)}$ be the stopping time of first transmission (starting from $E_0 = 0$), under $f^{(k)}$. Then

- $L_\beta^{(k)}(e) = (1-\beta)\mathbb{E}\left[\sum_{t=0}^{\tau^{(k)}-1} \beta^t d(E_t) \;\middle|\; E_0 = 0\right]$.
- $M_\beta^{(k)}(e) = (1-\beta)\mathbb{E}\left[\sum_{t=0}^{\tau^{(k)}-1} \beta^t \;\middle|\; E_0 = 0\right]$.

Regenerative process: The process $\{X_t\}_{t=0}^{\infty}$, if there exist $0 \leq T_0 < T_1 < T_2 < \cdots$ such that $\{X_t\}_{t=T_k+s}^{\infty}$, $s \geq 0$,

- has the same distribution as $\{X_t\}_{t=T_0+s}^{\infty}$,
- is independent of $\{X_t\}_{t=0}^{T_k}$.

Step 1: Main idea

Proof technique followed after Lerma, Lasserre - Discrete-time Markov control processes: basic optimality criteria, Springer

- The model satisfies certain assumptions (4.2.1, 4.2.2)
- Hence, the structural results extend to the infinite horizon discounted cost setup (Theorem 4.2.3)
- The discounted model satisfies some more assumptions (4.2.1, 5.4.1)
- Hence, structural results extend to long-term average setup (Theorem 5.4.3)

- Assumption 4.2.1 - The one-stage cost is l.s.c, non-negative and inf-compact on the set of feasible state-action pairs. The stochastic kernel ϕ is strongly continuous.
- Assumption 4.2.2 - There exists a strategy π such that the value function $V(\pi, x) < \infty$ for each state $x \in X$.
- Theorem 4.2.3 - Suppose Assumptions 4.2.1 and 4.2.2 hold. Then, in the discounted setup, there exists a selector which attains the minimum V_β^* and the optimal strategy, if it exists, is deterministic stationary.
- Assumption 5.4.1 - There exixts a state $z \in X$ and scalars $\alpha \in (0, 1)$ and $M \geq 0$ such that
  1. $(1 - \beta)V_\beta^*(z) \leq M$, $\forall \beta \in [\alpha, 1)$.
  2. Let $h_\beta(x) := V_\beta(x) - V_\beta(z)$. There exists $N \geq 0$ and a non-negative (not necessarily measurable) function $b(\cdot)$ on X such that $-N \leq h_\beta(x) \leq b(x)$, $\forall x \in X$ and $\beta \in [\alpha, 1)$.

- Theorem 5.4.3 - Suppose that Assumption 4.2.1 holds. Then the optimal stategy for average cost setup is deterministic stationary and is obtained by taking limit $\beta \uparrow 1$. The vanishing discount method is applicable and is employed to compute the optimal performance.

Step 1: Optimal threshold-type transmitter strategy for long-term average setup

The DP satisfies some suitable conditions so that, the vanishing discount approach is applicable.

Step 1: Optimal threshold-type transmitter strategy for long-term average setup

The DP satisfies some suitable conditions so that, the vanishing discount approach is applicable.

- For discounted setup, $\beta \in (0, 1]$, optimal transmitting strategy $f^*_\beta(\cdot; \lambda)$ is deterministic, threshold-type.
- Let $f^*(\cdot; \lambda)$ be any limit point of $f^*_\beta(\cdot; \lambda)$ as $\beta \uparrow 1$. Then the time-homogeneous transmission strategy $f^*(\cdot; \lambda)$ is optimal for $\beta = 1$ (the long-term average setup).
- Performance of optimal strategy: $C^*(\lambda) := C(f^*, g^*; \lambda) := \inf_{(f,g)} C(f, g; \lambda) = \lim_{\beta \uparrow 1} C^*_\beta(\lambda)$

Step 1: The SEN conditions

For any $\lambda \geq 0$, the value function $V_\beta(\cdot;\lambda)$, as given by a suitable DP, satisfies the following SEN conditions of [Lerma, Lasserre]:

SEN conditions

(S1) There exists a reference state $e_0 \in \mathbb{R}$ and a non-negative scalar M_λ such that $V_\beta(e_0,\lambda) < M_\lambda$ for all $\beta \in (0,1)$.

(S2) Define $h_\beta(e;\lambda) = (1-\beta)^{-1}[V_\beta(e;\lambda) - V_\beta(e_0;\lambda)]$. There exists a function $K_\lambda : \mathbb{Z} \to \mathbb{R}$ such that $h_\beta(e;\lambda) \leq K_\lambda(e)$ for all $e \in \mathbb{R}$ and $\beta \in (0,1)$.

(S3) There exists a non-negative (finite) constant L_λ such that $-L_\lambda \leq h_\beta(e;\lambda)$ for all $e \in \mathbb{R}$ and $\beta \in (0,1)$.

Step 2: Performance of threshold based strategies

Cost until first transmission: solution of FIE

Let $\tau^{(k)}$ be the stopping time when the Gauss-Markov process starting at state 0 at time $t = 0$ enters the set $\{e \in \mathbb{R} : |e| \geq \}$. Expected distortion incurred until stopping and expected stopping time under $f^{(k)}$ are solutions of Fredholm integral equations of second kind.

$$L^{(k)}(e) = e^2 + \int_{-k}^{k} \phi(w-e) L^{(k)}(w) dw;$$
$$M^{(k)}(e) = 1 + \int_{-k}^{k} \phi(w-e) M^{(k)}(w) dw.$$

Note that we have dropped the subscript 1 for ease of notation.

Step 2: Performance of threshold based strategies

Solutions to FIE

- Let $\mathcal{C}^{(k)}$ denote the space of bounded functions from $[-k, k]$ to $\mathbb{R}$. Define the operator $\mathcal{B}^{(k)} : \mathcal{C}^{(k)} \to \mathcal{C}^{(k)}$ as follows. For any $v \in \mathcal{C}^{(k)}$,

$$[\mathcal{B}^{(k)} v](e) = \int_{-k}^{k} \phi(w - e) v(w) dw.$$

- The operator $\mathcal{B}^{(k)}$ is a contraction
- Hence, FIE has a unique bounded solution $L^{(k)}$ and $M^{(k)}$.

Step 2: Performance of threshold based strategies

Renewal relationship

$$D^{(k)}(0) = \frac{L^{(k)}(0)}{M^{(k)}(0)}, \quad N^{(k)}(0) = \frac{1}{M^{(k)}(0)}$$

Step 2: Performance of threshold based strategies

Renewal relationship

$$D^{(k)}(0) = \frac{L^{(k)}(0)}{M^{(k)}(0)}, \quad N^{(k)}(0) = \frac{1}{M^{(k)}(0)}$$

Properties

- $L^{(k)}$ and $M^{(k)}$ are continuous, differentiable and monotonically increasing in k.
- $D^{(k)}(0)$ and $N^{(k)}(0)$ are continuous and differentiable in k. Furthermore, $N^{(k)}(0)$ is strictly decreasing in k.
- $D^{(k)}(0)$ is increasing in k.

Step 3: Identify critical Lagrange multipliers

Critical Lagrange multipliers

$$\lambda = -\frac{D_k^{(k)}(0)}{N_k^{(k)}(0)}, \tag{1}$$

Step 3: Identify critical Lagrange multipliers

Critical Lagrange multipliers

$$\lambda = -\frac{D_k^{(k)}(0)}{N_k^{(k)}(0)}, \tag{1}$$

Optimal transmission startegy

$(f^{(k)}, g^*)$ is $\lambda^{(k)}$-optimal for Lagrange relaxation. Furthermore, for any $k > 0$, there exists a $\lambda = \lambda^{(k)} \geq 0$ that satisfies (1).

Step 3: Identify critical Lagrange multipliers

Critical Lagrange multipliers

$$\lambda = -\frac{D_k^{(k)}(0)}{N_k^{(k)}(0)}, \quad (1)$$

Optimal transmission startegy

$(f^{(k)}, g^*)$ is $\lambda^{(k)}$-optimal for Lagrange relaxation. Furthermore, for any $k > 0$, there exists a $\lambda = \lambda^{(k)} \geq 0$ that satisfies (1).

Proof

- The choice of λ implies that $C_k^{(k)}(0; \lambda) = 0$. Hence strategy $(f^{(k)}, g^*)$ is λ-optimal.
- $\lambda^{(k)} \geq 0$, by the properties of $D^{(k)}(0)$ and $N^{(k)}(0)$.

Step 4: The constrained setup

A strategy (f°, g°) is optimal for a constrained optimization problem, if

Sufficient conditions for optimality [Sennott, 1999]

(C1) $N(f^\circ, g^\circ) = \alpha$,

(C2) There exists a Lagrange multiplier $\lambda^\circ \geq 0$ such that (f°, g°) is optimal for $C(f, g; \lambda^\circ)$.

Step 4: The constrained setup

- For $\alpha \in (0, 1)$, let $k^*(\alpha)$ be such that $N^{(k^*(\alpha))} = \alpha$. Find $k^*(\alpha)$ for a given α;

Optimal deterministic strategy $f^* = f^{(k^*(\alpha))}$.

Step 4: The constrained setup

- For $\alpha \in (0,1)$, let $k^*(\alpha)$ be such that $N^{(k^*(\alpha))} = \alpha$. Find $k^*(\alpha)$ for a given α;

Optimal deterministic strategy $f^* = f^{(k^*(\alpha))}$.

Proof

- (C1) is satisfied by $f^\circ = f^{(k^*(\alpha))}$ and $g^\circ = g^*$.
- For $k^*(\alpha)$, we can find a λ satisfying (1). Hence we have that $(f^{(k^*(\alpha))}, g^*)$ is optimal for $C(f, g; \lambda)$.
- Thus, $(f^{(k^*(\alpha))}, g^*)$ satisfies (C2).
- $D^*(\alpha) := D(f^{(k^*(\alpha))}, g^*) = D^{(k^*(\alpha))}(0)$

Algorithm

Algorithm 1: Computation of $D^*_\beta(\alpha)$

input : $\alpha \in (0,1)$, $\beta \in (0,1]$, $\varepsilon \in \mathbb{R}_{>0}$

output: $D_\beta^{(k^\circ)}(\alpha)$, where $|N_\beta^{(k^\circ)}(0) - \alpha| < \varepsilon$

Pick $\underline{k}$ and $\bar{k}$ such that $N_\beta^{(\underline{k})}(0) < \alpha < N_\beta^{(\bar{k})}(0)$

$k^\circ = (\underline{k} + \bar{k})/2$

while $|N_\beta^{(k^\circ)}(0) - \alpha| > \varepsilon$ **do**

    if $N_\beta^{(k^\circ)}(0) < \alpha$ **then**

        $\underline{k} = k^\circ$

    else

        $\bar{k} = k^\circ$

    $k^\circ = (\underline{k} + \bar{k})/2$

return $D_\beta^{(k^\circ)}(\alpha)$