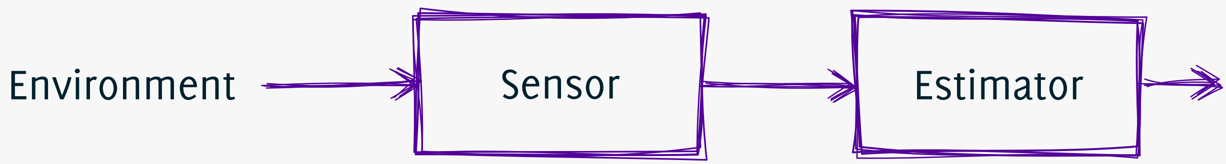
Estimation with active sensing

Aditya Mahajan
McGill University

February 11, 2011. ITA Workshop

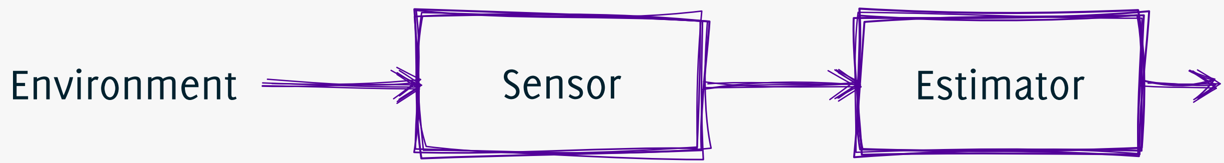
System Model



Sensor transmits packet to the estimator

- Each transmission incurs a constant cost
  (energy required to switch on the radio and transmit a packet).
- Trade-off estimation quality with transmission energy
- The sensor decides when and how to transmit
  (as opposed to the estimator scheduling the transmissions
  . . . or the sensor communicating without coding)

System Model

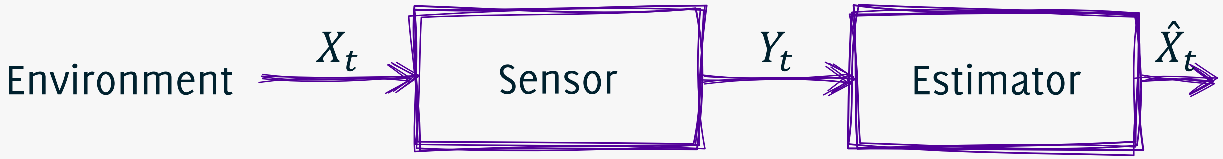


Sensor transmits packet to the estimator

- Each transmission incurs a constant cost
  (energy required to switch on the radio and transmit a packet).
- Trade-off estimation quality with transmission energy
- The sensor decides when and how to transmit
  (as opposed to the estimator scheduling the transmissions
  . . . or the sensor communicating without coding)

When to communicate? And how to communicate?

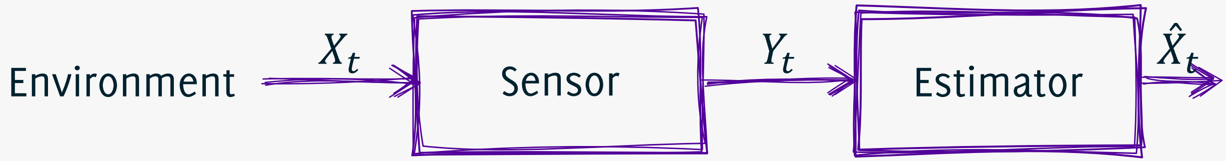
System Model



- Environment

  First order Markov process $\{X_t, t = 1, 2, \ldots\}$, $X_t \in \mathbb{X}$

System Model



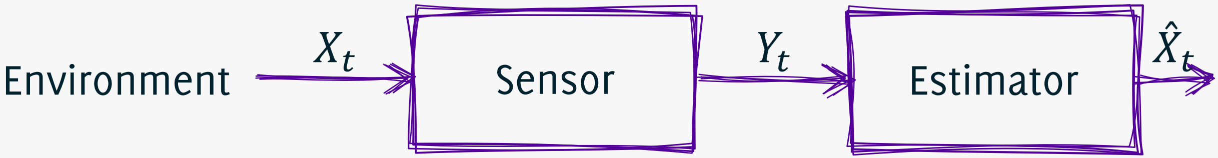
- Environment

  First order Markov process $\{X_t, t = 1, 2, \ldots\}$, $X_t \in \mathbb{X}$

- Sensor

  $$Y_t = f_t(X_{1:t}, Y_{1:t-1}), \quad Y_t \in \mathbb{Y} := \mathbb{X} \cup \{\mathbb{b}\}$$

System Model



- Environment

First order Markov process $\{X_t, t = 1, 2, \ldots\}$, $X_t \in \mathbb{X}$

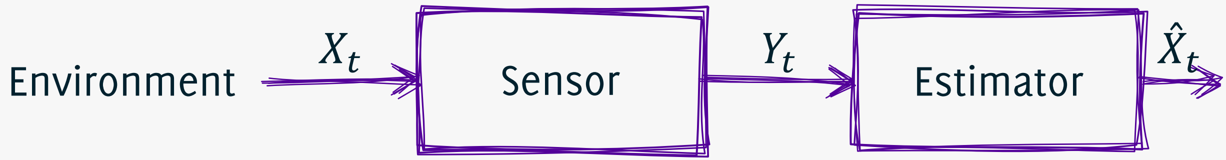
- Sensor

$$Y_t = f_t(X_{1:t}, Y_{1:t-1}), \quad Y_t \in \mathbb{Y} := \mathbb{X} \cup \{\mathbb{b}\}$$

- Estimator

$$\hat{X}_t = g_t(Y_{1:t}), \quad \hat{X}_t \in \mathbb{X}$$

System Model



- Environment

  First order Markov process $\{X_t, t = 1, 2, \ldots\}$, $X_t \in \mathbb{X}$

- Sensor

$$Y_t = f_t(X_{1:t}, Y_{1:t-1}), \quad Y_t \in \mathbb{Y} := \mathbb{X} \cup \{\mathbb{b}\}$$

- Estimator

$$\hat{X}_t = g_t(Y_{1:t}), \quad \hat{X}_t \in \mathbb{X}$$

- Cost $c(Y_t) + d(X_t, \hat{X}_t)$

  - $c(y) = \begin{cases} 0, & y = \mathbb{b}, \\ c^*, & y \in \mathbb{X} \end{cases}$
  - d is a metric on $\mathbb{X}$

Problem Formulation

- Transmission policy and estimation policies

$$\boldsymbol{f} = (f_1, f_2, \dots, f_T) \qquad \boldsymbol{g} = (g_1, g_2, \dots, g_T)$$

- Performance of a policy

$$\mathcal{J}(\boldsymbol{f}, \boldsymbol{g}) \coloneqq \mathbb{E}^{\boldsymbol{f},\boldsymbol{g}}\Big[\sum_{t=1}^{T} c(Y_t) + d(X_t, \hat{X}_t)\Big]$$

- Causal Policies $(\mathcal{F}, \mathcal{G})$

$$\mathcal{F} = \prod_{t=1}^{T} F_t, \qquad F_t = \big(\mathbb{X}^t \times \mathbb{Y}^{t-1} \mapsto \mathbb{Y}\big)$$

$$\mathcal{G} = \prod_{t=1}^{T} G_t, \qquad G_t = \big(\mathbb{Y}^t \mapsto \mathbb{X}\big)$$

Problem Formulation

- Transmission policy and estimation policies

$$\boldsymbol{f} = (f_1, f_2, \dots, f_T) \qquad \boldsymbol{g} = (g_1, g_2, \dots, g_T)$$

- Performance of a policy

$$\mathcal{J}(\boldsymbol{f}, \boldsymbol{g}) := \mathbb{E}^{\boldsymbol{f},\boldsymbol{g}}\Big[\sum_{t=1}^{T} c(Y_t) + d(X_t, \hat{X}_t)\Big]$$

- Causal Policies $(\mathcal{F}, \mathcal{G})$

$$\mathcal{F} = \prod_{t=1}^{T} F_t, \qquad F_t = \big(\mathbb{X}^t \times \mathbb{Y}^{t-1} \mapsto \mathbb{Y}\big)$$

$$\mathcal{G} = \prod_{t=1}^{T} G_t, \qquad G_t = \big(\mathbb{Y}^t \mapsto \mathbb{X}\big)$$

P1: Choose $(\boldsymbol{f}, \boldsymbol{g}) \in (\mathcal{F}, \mathcal{G})$ to minimize $\mathcal{J}(\boldsymbol{f}, \boldsymbol{g})$

Salient Features

- Sequential dynamic team

  Sensor and estimator have access to different information at different times.

Salient Features

- Sequential dynamic team

  Sensor and estimator have access to different information at different times.

- Causal real-time communication

  Equivalent to a real-time communication system with input cost

Literature overview

- Huge literature on sensor sleep scheduling and sensor “censoring”.

Literature overview

- Huge literature on sensor sleep scheduling and sensor "censoring".
- Most relevant results: Imer and Basar, 2005, Lipsa and Martins, 2007

  Consider $\{X_t, t = 1, 2, \ldots\}$ is a Gauss-Markov process

Literature overview

- Huge literature on sensor sleep scheduling and sensor "censoring".
- Most relevant results: Imer and Basar, 2005, Lipsa and Martins, 2007

  Consider $\{X_t, t = 1, 2, \ldots\}$ is a Gauss-Markov process
- Structure of optimal policy: Communicate if

$$|X_t - \text{prediction if sensor does not transmit}| \geq \tau_t$$

Literature overview

- Huge literature on sensor sleep scheduling and sensor "censoring".
- Most relevant results: Imer and Basar, 2005, Lipsa and Martins, 2007

  Consider $\{X_t, t = 1, 2, \ldots\}$ is a Gauss-Markov process
- Structure of optimal policy: Communicate if

$$|X_t - \text{prediction if sensor does not transmit}| \geq \tau_t$$

- Implication: Optimal policy is easy to implement . . .

  . . . but the intuition relies of the Markov process being Gaussian.

For non-Gaussian processes, is the optimal policy easy to implement?

Solution approach

- Identify irrelevant information at sensor

  Fix estimator. Set up the problem at the sensor as a MDP

- Find structure of optimal transmission and estimation policy: Coding does not improve performance

  A sequence of interchange arguments

- Identify a dynamic programming decomposition

  Dynamic team problem. Use the coordinator approach proposed by [Mahajan, Nayyar, Teneketzis, 2008]

- Simplify DP using structural result

- Implementation issues

Removing irrelevant information

- Simple Causal policies $(\mathcal{F}', \mathcal{G})$

$$\mathcal{F}' = \prod_{t=1}^{T} F'_t,$$

Removing irrelevant information

- Simple Causal policies $(\mathcal{F}', \mathcal{G})$

$$\mathcal{F}' = \prod_{t=1}^{T} F'_t,$$

$$F'_t = \Big\{ f_t \in F_t : \forall x^{(1)}_{1:t-1}, x^{(2)}_{1:t-1} \in \mathbb{X}^{t-1}, \forall x_t \in \mathbb{X}, \forall y_{1:t-1} \in \mathbb{Y},$$
$$\text{we have } f_t(x^{(1)}_{1:t-1}, x_t, y_{1:t-1}) = f_t(x^{(2)}_{1:t-1}, x_t, y_{1:t-1}) \Big\}.$$

Removing irrelevant information

- Simple Causal policies $(\mathcal{F}', \mathcal{G})$

$$\mathcal{F}' = \prod_{t=1}^{T} F'_t,$$

$$F'_t = \Big\{ f_t \in F_t : \forall x^{(1)}_{1:t-1}, x^{(2)}_{1:t-1} \in \mathbb{X}^{t-1}, \forall x_t \in \mathbb{X}, \forall y_{1:t-1} \in \mathbb{Y},$$
$$\text{we have } f_t(x^{(1)}_{1:t-1}, x_t, y_{1:t-1}) = f_t(x^{(2)}_{1:t-1}, x_t, y_{1:t-1}) \Big\}.$$

or equivalently if $f_t \in F'_t$ then $y_t = f_t(x_t, y_{1:t-1})$

Removing irrelevant information

- Simple Causal policies $(\mathcal{F}', \mathcal{G})$

$$\mathcal{F}' = \prod_{t=1}^{T} F'_t,$$
$$F'_t = \Big\{ f_t \in F_t : \forall x^{(1)}_{1:t-1}, x^{(2)}_{1:t-1} \in \mathbb{X}^{t-1}, \forall x_t \in \mathbb{X}, \forall y_{1:t-1} \in \mathbb{Y}, \\ \text{we have } f_t(x^{(1)}_{1:t-1}, x_t, y_{1:t-1}) = f_t(x^{(2)}_{1:t-1}, x_t, y_{1:t-1}) \Big\}.$$

or equivalently if $f_t \in F'_t$ then $y_t = f_t(x_t, y_{1:t-1})$

- Proposition

There is no loss of optimality in restricting attention to simple causal policies.

Proof Outline

Fix an arbitrary estimation policy. Look at the problem of choosing the best transmission policy.

Define $Z_t = (X_t, Y_{1:t-1})$. We can show that

- $\mathbb{P}(Z_{t+1} | Z_{1:t}, Y_{1:t}) = \mathbb{P}(Z_{t+1} | Z_t, Y_t)$
- $\mathbb{E}[c(Y_t) + d(X_t, \hat{X}_t) | Z_{1:t}, Y_{1:t}] = \mathbb{E}[c(Y_t) + d(X_t, \hat{X}_t) | Z_t, Y_t]$

Proof Outline

Fix an arbitrary estimation policy. Look at the problem of choosing the best transmission policy.

Define $Z_t = (X_t, Y_{1:t-1})$. We can show that

- $\mathbb{P}(Z_{t+1} | Z_{1:t}, Y_{1:t}) = \mathbb{P}(Z_{t+1} | Z_t, Y_t)$
- $\mathbb{E}[c(Y_t) + d(X_t, \hat{X}_t) | Z_{1:t}, Y_{1:t}] = \mathbb{E}[c(Y_t) + d(X_t, \hat{X}_t) | Z_t, Y_t]$

Hence, $\{Z_t, t = 1, 2, \dots\}$ is a controlled Markov process. Thus, there is no loss of optimality in restricting attention to

$$Y_t = f_t(Z_t) = f_t(X_t, Y_{1:t-1})$$

Structure of optimal policies

- Structured Policies $(\mathcal{F}^*, \mathcal{G}^*)$

$$\mathcal{F}^\star = \prod_{t=1}^{T} F_t^\star,$$
$$F_t^\star = \Big\{ f_t \in F_t' : \forall x_t \in \mathbb{X},\ \forall y_{1:t-1} \in \mathbb{Y},\ f_t(x_t, y_{1:t-1}) \in \{x_t, \mathbb{b}\} \Big\}$$

Structure of optimal policies

- Structured Policies $(\mathcal{F}^*, \mathcal{G}^*)$

$$\mathcal{F}^\star = \prod_{t=1}^{T} F_t^\star,$$

$$F_t^\star = \Big\{ f_t \in F_t' : \forall x_t \in \mathbb{X},\ \forall y_{1:t-1} \in \mathbb{Y},\ f_t(x_t, y_{1:t-1}) \in \{x_t, \mathbb{b}\} \Big\}$$

and

$$\mathcal{G}^\star = \prod_{t=1}^{T} G_t^\star,$$

$$G_t^\star = \Big\{ g_t \in G_t : \forall y_{1:t-1} \in \mathbb{Y} \text{ and } \forall y_t \in \mathbb{X},\ g_t(y_{1:t}) = y_t \Big\}$$

Structure of optimal policies

- Structured Policies $(\mathcal{F}^*, \mathcal{G}^*)$

$$\mathcal{F}^\star = \prod_{t=1}^{T} F_t^\star,$$
$$F_t^\star = \Big\{ f_t \in F_t' : \forall x_t \in \mathbb{X}, \forall y_{1:t-1} \in \mathbb{Y}, f_t(x_t, y_{1:t-1}) \in \{x_t, \mathbb{b}\} \Big\}$$

and

$$\mathcal{G}^\star = \prod_{t=1}^{T} G_t^\star,$$
$$G_t^\star = \Big\{ g_t \in G_t : \forall y_{1:t-1} \in \mathbb{Y} \text{ and } \forall y_t \in \mathbb{X}, g_t(y_{1:t}) = y_t \Big\}$$

Theorem: There is no loss of optimality in restricting attention to structured policies.

Implication: Real-time coding does not improve performance

Proof Outline

- Define $(\mathcal{F}^{(s)}, \mathcal{G}^{(s)})$, $s = 0, 1, \ldots, T$ as follows

$$(F_t^{(s)}, G_t^{(s)}) = \begin{cases} (F_t', G_t) & t < s \\ (F_t^\star, G_t^\star) & t \geq s \end{cases}$$

Proof Outline

- Define $(\mathcal{F}^{(s)}, \mathcal{G}^{(s)})$, $s = 0, 1, \ldots, T$ as follows

$$(F_t^{(s)}, G_t^{(s)}) = \begin{cases} (F_t', G_t) & t < s \\ (F_t^\star, G_t^\star) & t \geq s \end{cases}$$

- Lemma

For any $(\boldsymbol{f}^{(s+1)}, \boldsymbol{g}^{(s+1)}) \in (\mathcal{F}^{(s+1)}, \mathcal{G}^{(s+1)})$, there is $(\boldsymbol{f}^{(s)}, \boldsymbol{g}^{(s)}) \in (\mathcal{F}^{(s)}, \mathcal{G}^{(s)})$, such that

$$\mathcal{J}(\boldsymbol{f}^{(s)}, \boldsymbol{g}^{(s)}) \leq \mathcal{J}(\boldsymbol{f}^{(s+1)}, \boldsymbol{g}^{(s+1)})$$

Proof Outline

- Define $(\mathcal{F}^{(s)}, \mathcal{G}^{(s)})$, $s = 0, 1, \ldots, T$ as follows

$$(F_t^{(s)}, G_t^{(s)}) = \begin{cases} (F_t', G_t) & t < s \\ (F_t^\star, G_t^\star) & t \geq s \end{cases}$$

- Lemma

For any $(\boldsymbol{f}^{(s+1)}, \boldsymbol{g}^{(s+1)}) \in (\mathcal{F}^{(s+1)}, \mathcal{G}^{(s+1)})$, there is $(\boldsymbol{f}^{(s)}, \boldsymbol{g}^{(s)}) \in (\mathcal{F}^{(s)}, \mathcal{G}^{(s)})$, such that

$$\mathcal{J}(\boldsymbol{f}^{(s)}, \boldsymbol{g}^{(s)}) \leq \mathcal{J}(\boldsymbol{f}^{(s+1)}, \boldsymbol{g}^{(s+1)})$$

- Lemma $\implies$ Theorem.

$$(\mathcal{F}^{(T+1)}, \mathcal{G}^{(T+1)}) = (\mathcal{F}', \mathcal{G}), \qquad (\mathcal{F}^{(0)}, \mathcal{G}^{(0)}) = (\mathcal{F}^\star, \mathcal{G}^\star)$$

So how do we find an optimal policy?

Dynamic programming decomposition

- Non-classical information structure

$$Y_t = f_t(X_t, Y_{1:t-1}) \qquad \hat{X}_t = g_t(Y_{1:t-1}, Y_t)$$

Sensor | Estimator

Dynamic programming decomposition

- Non-classical information structure

$$Y_t = f_t(X_t, Y_{1:t-1}) \qquad \hat{X}_t = g_t(Y_{1:t-1}, Y_t)$$

Sensor Estimator

- Belongs to the class of tractable non-classical information structures identified in [Mahajan, Nayyar, Teneketzis, 2008]

Dynamic programming decomposition

- Non-classical information structure

$$Y_t = f_t(X_t, Y_{1:t-1}) \qquad \hat{X}_t = g_t(Y_{1:t-1}, Y_t)$$

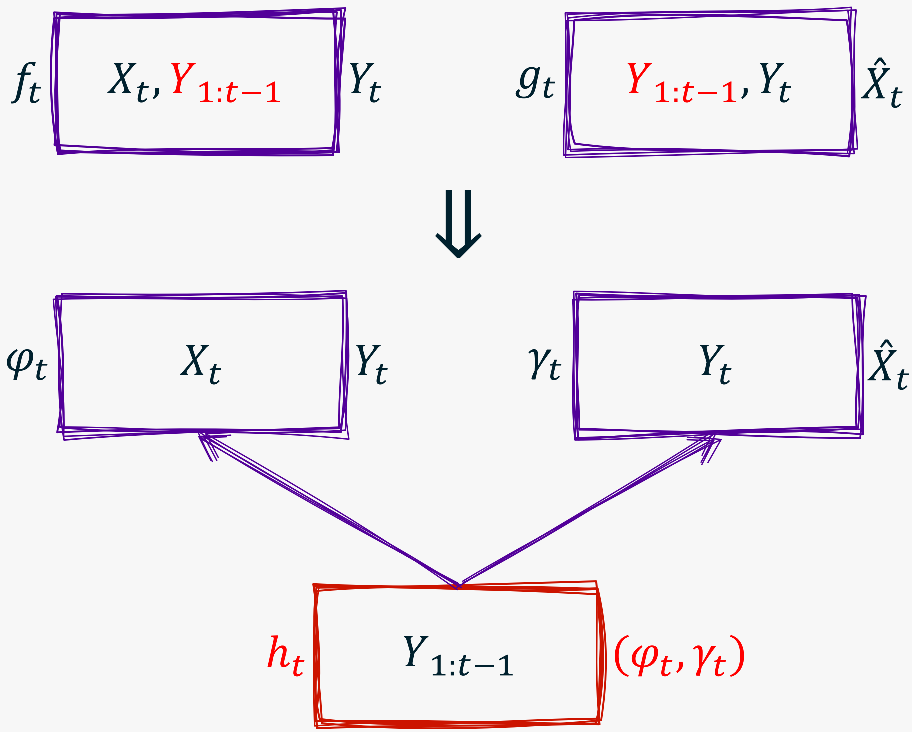
Sensor — Estimator

- Belongs to the class of tractable non-classical information structures identified in [Mahajan, Nayyar, Teneketzis, 2008]
- Look at the problem from the p.o.v. of a coordinator that observes the common data between the sensor and the estimator.
- The local data does not increase with time.

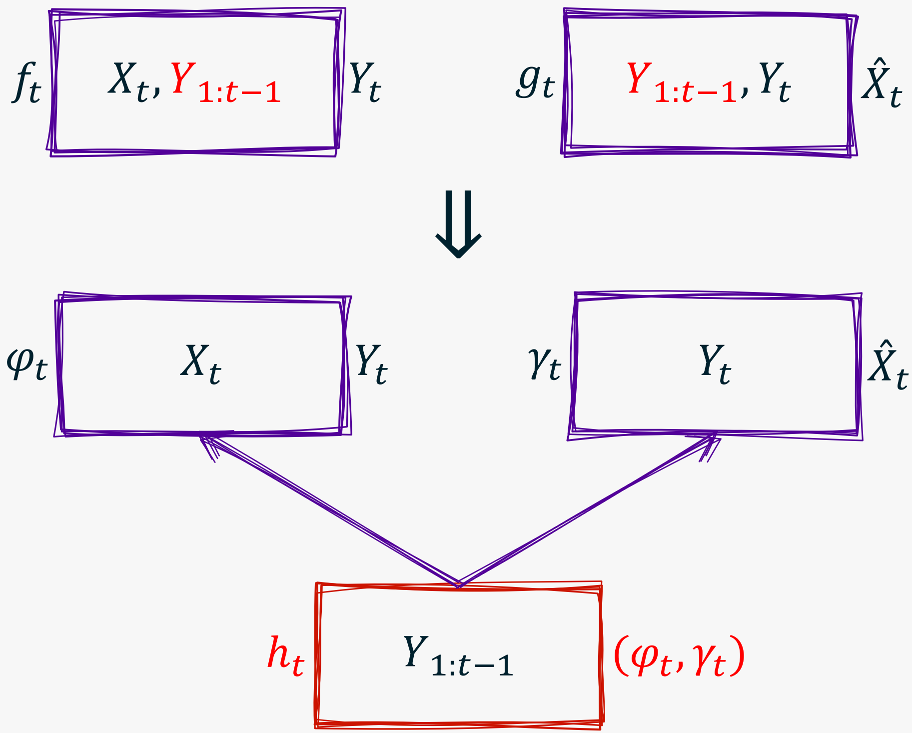
The coordinated system

$$f_t \; \boxed{X_t, Y_{1:t-1}} \; Y_t \qquad\qquad g_t \; \boxed{Y_{1:t-1}, Y_t} \; \hat{X}_t$$

The coordinated system



The coordinated system



Both systems are equivalent

The dynamic programming decomposition

The coordinated system is a centralized (single agent) POMDP

The dynamic programming decomposition

The coordinated system is a centralized (single agent) POMDP

Structure Result

Let $\pi_t = \mathbb{P}(X_t | Y_{1:t-1}, \varphi_{1:t-1}, \gamma_{1:t-1})$.

$$(\varphi_t, \gamma_t) = h_t(\pi_t)$$

The dynamic programming decomposition

The coordinated system is a centralized (single agent) POMDP

- **Structure Result**

  Let $\pi_t = \mathbb{P}(X_t | Y_{1:t-1}, \varphi_{1:t-1}, \gamma_{1:t-1})$.

$$(\varphi_t, \gamma_t) = h_t(\pi_t)$$

- **Dynamic Program**

$$V_t(\pi_t) = \min_{\varphi_t, \gamma_t} \left\{ \mathbb{E}\Big[c(X_t) + d(X_t, \hat{X}_t) + V_{t+1}(\text{update}(\pi_t, Y_t)) \,\Big|\, \pi_t, \varphi_t, \gamma_t \Big] \right\}$$

The dynamic programming decomposition

The coordinated system is a centralized (single agent) POMDP

- **Structure Result**

  Let $\pi_t = \mathbb{P}(X_t | Y_{1:t-1}, \varphi_{1:t-1}, \gamma_{1:t-1})$.

  $$(\varphi_t, \gamma_t) = h_t(\pi_t)$$

- **Dynamic Program**

  $$V_t(\pi_t) = \min_{\varphi_t, \gamma_t} \left\{ \mathbb{E}\left[c(X_t) + d(X_t, \hat{X}_t) + V_{t+1}(\text{update}(\pi_t, Y_t)) \,\middle|\, \pi_t, \varphi_t, \gamma_t \right] \right\}$$

- Each step is a functional optimization problem

Refinement of DP

Restrict attention to structured policies

Refinement of DP

Restrict attention to structured policies

- $\forall f_t \in F_t^\star, \quad f_t(x_t, y_{1:t-1}) \in \{x_t, \mathbb{b}\}.$

  Thus, $\varphi_t(\cdot) \equiv \left\{x \in \mathbb{X} : \varphi_t(x_t) = x_t\right\} =: C_t$ (communication set)

Refinement of DP

Restrict attention to structured policies

- $\forall f_t \in F_t^\star, \quad f_t(x_t, y_{1:t-1}) \in \{x_t, \mathbb{b}\}.$

  Thus, $\varphi_t(\cdot) \equiv \big\{x \in \mathbb{X} : \varphi_t(x_t) = x_t\big\} =: C_t$ (communication set)

- $\forall g_t \in G_t^\star, \quad g_t(y_{1:t-1}, y_t) = \begin{cases} y_t & \text{if } y_t \in \mathbb{X}, \\ \hat{x}_t^\star, & \text{if } y_t = \mathbb{b} \end{cases}$

  Thus, $\gamma_t(\cdot) \equiv \hat{x}_t^\star$

Refinement of DP

Restrict attention to structured policies

- $\forall f_t \in F_t^\star, \quad f_t(x_t, y_{1:t-1}) \in \{x_t, 𝕓\}.$

  Thus, $\varphi_t(\cdot) \equiv \big\{x \in \mathbb{X} : \varphi_t(x_t) = x_t\big\} =: C_t$ (communication set)

- $\forall g_t \in G_t^\star, \quad g_t(y_{1:t-1}, y_t) = \begin{cases} y_t & \text{if } y_t \in \mathbb{X}, \\ \hat{x}_t^\star, & \text{if } y_t = 𝕓 \end{cases}$

  Thus, $\gamma_t(\cdot) \equiv \hat{x}_t^\star$

- Dynamic Program

$$V_t(\pi_t) = \min_{C_t, \hat{x}_t^\star} \Big\{ \mathbb{E}\big[c(X_t) + d(X_t, \hat{X}_t) + V_{t+1}(\text{update}(\pi_t, Y_t)) \,\big|\, \pi_t, \varphi_t, \gamma_t\big] \Big\}$$

Refinement of DP

Restrict attention to structured policies

- $\forall f_t \in F_t^\star, \quad f_t(x_t, y_{1:t-1}) \in \{x_t, \mathbb{b}\}.$

  Thus, $\varphi_t(\cdot) \equiv \big\{x \in \mathbb{X} : \varphi_t(x_t) = x_t\big\} =: C_t$ (communication set)

- $\forall g_t \in G_t^\star, \quad g_t(y_{1:t-1}, y_t) = \begin{cases} y_t & \text{if } y_t \in \mathbb{X}, \\ \hat{x}_t^\star, & \text{if } y_t = \mathbb{b} \end{cases}$

  Thus, $\gamma_t(\cdot) \equiv \hat{x}_t^\star$

- Dynamic Program

$$V_t(\pi_t) = \min_{C_t, \hat{x}_t^\star} \Big\{ \mathbb{E}\big[c(X_t) + d(X_t, \hat{X}_t) + V_{t+1}(\text{update}(\pi_t, Y_t)) \,\big|\, \pi_t, \varphi_t, \gamma_t \big] \Big\}$$

- Can be converted to countable state MDP.

Implementation issues

For a fixed transmission policy, $\pi_t \equiv (U_t, N_t)$ where

- U_t last non-blank transmission
- N time since last non-blank transmission

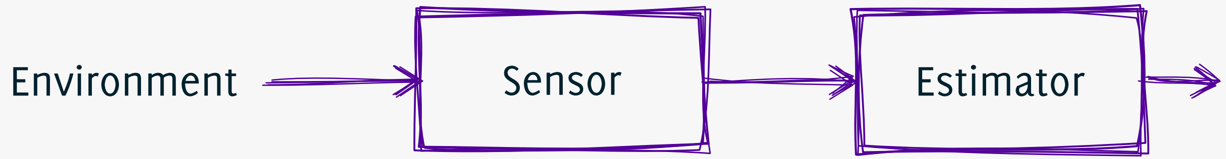
Implementation issues

For a fixed transmission policy, $\pi_t \equiv (U_t, N_t)$ where

- U_t last non-blank transmission
- N time since last non-blank transmission

- Sensor: $Y_t = f_t(X_t, U_t, N_t)$
- Estimator: $\hat{X}_t = g_t(Y_t, U_t, N_t)$

Conclusion



- Use ideas from dynamic team theory and real-time communications
  - remove irrelevant data at sensor
  - coding does not improve performance
  - DP decomposition by investigating the coordinated system
  - Simplify DP using structure of optimal policy
- Transformed to denumerable MDP with set valued actions

$$V_t(\pi_t) = \min_{C_t, \hat{x}_t^\star} \left\{ \mathbb{E}\Big[c(X_t) + d(X_t, \hat{X}_t) + V_{t+1}(\text{update}(\pi_t, Y_t)) \,\Big|\, \pi_t, \varphi_t, \gamma_t \Big] \right\}$$

Unanswered questions

- Monotonicity of communication sets

  Is $C_t(\pi) \subseteq C_t(\pi \cdot P)$?

Unanswered questions

- Monotonicity of communication sets

  Is $C_t(\pi) \subseteq C_t(\pi \cdot P)$?

- Weakest sufficient conditions to get threshold-type transmission policy?

  Or, equivalently $\exists\, \text{index}_t : \mathbb{X} \mapsto \mathbb{R} \ni C_t = \left\{x \in \mathbb{X} : \text{index}_t(x) > \tau_t\right\}$

Unanswered questions

- Monotonicity of communication sets

  Is $C_t(\pi) \subseteq C_t(\pi \cdot P)$?

- Weakest sufficient conditions to get threshold-type transmission policy?

  Or, equivalently $\exists\, \mathrm{index}_t : \mathbb{X} \mapsto \mathbb{R} \ni C_t = \left\{x \in \mathbb{X} : \mathrm{index}_t(x) > \tau_t\right\}$

- Why is this talk in the middle of two talks on bandits?

Unanswered questions

- Monotonicity of communication sets

  Is $C_t(\pi) \subseteq C_t(\pi \cdot P)$?

- Weakest sufficient conditions to get threshold-type transmission policy?

  Or, equivalently $\exists\, \mathrm{index}_t : \mathbb{X} \mapsto \mathbb{R} \ni C_t = \left\{x \in \mathbb{X} : \mathrm{index}_t(x) > \tau_t\right\}$

- Why is this talk in the middle of two talks on bandits?

  Consider "symmetric" Markov chains and a fixed decoding rule. Finding the optimal transmission policy is tangentially related to [Katehakis and Veinott, 1987] characterization of the Gittin's index as a restart in i problem.

  Could we reduce the problem to a 1-1/2 armed bandit problem?

Thank you.

Questions?