Partial conservation laws and indexability: past, present, and future

José Niño-Mora
http://alum.mit.edu/www/jnimora

Department of Statistics

Universidad
Carlos III de Madrid
www.uc3m.es

Keynote talk at *Workshop on restless bandits, index policies and applications in reinforcement learning*
Grenoble, 20–21 November, 2023

Work partially supported by grant PID2019-109196GB-I00 funded by MCIN/AEI/10.13039/501100011033

MDP model: optimal control of resource-fueled project

- Discrete-time **restless bandit**, modeling a resource-fueled **project**:
  - States (finite): $X(t) \in \mathcal{N} \triangleq \{1, 2, \ldots, n\}$
  - Actions (binary): $A(t) \in \mathcal{A} \triangleq \{0, 1\}$
  - One-period rewards: r_i^a
  - One-period quantity of resource consumed: $q_i^a \quad (0 < q_i^1 > q_i^0 \geqslant 0)$
  - Transition probabilities: p_{ij}^a
  - Discount factor: $0 < \beta < 1$
  - Π: admissible policies (stationary is enough)
- For each **resource price** $\lambda \in \mathbb{R}$, consider **λ-price problem**:

$$\underset{\pi \in \Pi}{\text{maximize}}\ \mathbb{E}_i^\pi \left[\sum_{t=0}^{\infty} \left(r_{X(t)}^{A(t)} - \lambda q_{X(t)}^{A(t)} \right) \beta^t \right]$$

- **Optimal value function**: $V_i^*(\lambda)$

λ-price problem & Bellman equations

- For each **resource price** $\lambda \in \mathbb{R}$, consider **λ-price problem**:

$$\underset{\pi \in \Pi}{\text{maximize}}\ \mathbb{E}_i^\pi \left[\sum_{t=0}^{\infty} \big(r_{X(t)}^{A(t)} - \lambda q_{X(t)}^{A(t)} \big) \beta^t \right]$$

- **Opt. val. funct.** $V_i^*(\lambda)$ & opt. policies through **Bellman equations**:

$$V_i^*(\lambda) = \max_{a \in \{0,1\}} r_i^a - \lambda q_i^a + \beta \sum_{j \in \mathcal{N}} p_{ij}^a V_j^*(\lambda), \quad i \in \mathcal{N}$$

- Write them as:

$$V_i^*(\lambda) = \max_{a \in \{0,1\}} V_i^{\langle a,* \rangle}(\lambda), \quad i \in \mathcal{N}$$

Indexability and Whittle index

- $V_i^*(\lambda)$ and optimal policies determined by **Bellman equations**:

$$V_i^*(\lambda) = \max_{a \in \{0,1\}} V_i^{\langle a,* \rangle}(\lambda), \quad i \in \mathcal{N}$$

- Consider the marginal value function $v_i^*(\lambda) \triangleq V_i^{\langle 1,* \rangle}(\lambda) - V_i^{\langle 0,* \rangle}(\lambda)$

Call the project indexable if, for each state i:

1. The eqn. $v_i^*(\lambda) = 0$ has a unique root $\lambda = \lambda_i^*$ (Whittle index)
2. $v_i^*(\lambda) > 0$ for $\lambda < \lambda_i^*$
3. $v_i^*(\lambda) < 0$ for $\lambda > \lambda_i^*$

The submodularity approach to indexability

- $V_i^*(\lambda)$ and optimal policies determined by **Bellman equations**:

$$V_i^*(\lambda) = \max_{a \in \{0,1\}} V_i^{\langle a,*\rangle}(\lambda), \quad i \in \mathcal{N}$$

- Write $v_i^*(\lambda) \triangleq V_i^{\langle 1,*\rangle}(\lambda) - V_i^{\langle 0,*\rangle}(\lambda)$
- Suppose can prove that $V_i^{\langle a,*\rangle}(\lambda)$ is strictly supermodular in (i, a)
- i.e., $v_i^*(\lambda)$ is increasing in i (for state ordering $i = 1, \ldots, n$):

$$v_1^*(\lambda) < v_2^*(\lambda) < \cdots < v_n^*(\lambda)$$

- This implies optimality of threshold policies

The submodularity approach to indexability

- Suppose one can prove that $\forall \lambda$, $v_i^*(\lambda)$ is increasing in $i = 1, \ldots, n$:

$$v_1^*(\lambda) < v_2^*(\lambda) < \cdots < v_n^*(\lambda)$$

- This implies optimality of threshold policies

> - z-policy: has active set $S = \{j \in \mathcal{N} \colon j > z\}$
> - $\lambda \in \Lambda_0 \triangleq \{\lambda \colon 0 \leqslant v_1^*(\lambda)\}$: 0-policy optimal ($z^*(\lambda) = 0$)
> - $\lambda \in \Lambda_i \triangleq \{\lambda \colon v_i^*(\lambda) < 0 \leqslant v_{i+1}^*(\lambda)\}$: i-policy optimal ($z^*(\lambda) = i$)
> - $\lambda \in \Lambda_n \triangleq \{\lambda \colon v_n^*(\lambda) < 0\}$: n-policy optimal ($z^*(\lambda) = n$)

- For indexability: prove $\Lambda_0 = (-\infty, \lambda_1^*]$, $\Lambda_i = (\lambda_i^*, \lambda_{i+1}^*]$, $\Lambda_n = (\lambda_n^*, \infty)$
- Equivalently, need to further prove that $z^*(\lambda) \nearrow$, spanning $\{0, \ldots, n\}$

Reformulating thru project performance metrics

- **Reward metric:** $F_i^\pi \triangleq \mathbb{E}_i^\pi \left[\sum_{t=0}^{\infty} r_{X(t)}^{A(t)} \beta^t \right]$
- **Resource (usage) metric:** $G_i^\pi \triangleq \mathbb{E}_i^\pi \left[\sum_{t=0}^{\infty} q_{X(t)}^{A(t)} \beta^t \right]$
- **λ-price problem:** $\underset{\pi \in \Pi}{\text{maximize}}\ F_i^\pi - \lambda G_i^\pi$
- Stationary policies are enough: S-active policy, for $S \in 2^{\mathcal{N}}$
- **λ-price problem:** $\underset{S \in 2^{\mathcal{N}}}{\text{maximize}}\ F_i^S - \lambda G_i^S$
- **Optimal value function:** $V_i^*(\lambda) = \max_{S \in 2^{\mathcal{N}}} F_i^S - \lambda G_i^S$

Marginal project performance metrics

- **Marginal reward metric:** $f_i^S \triangleq F_i^{\langle 1,S \rangle} - F_i^{\langle 0,S \rangle}$

- **Marginal resource (usage) metric:** $g_i^S \triangleq G_i^{\langle 1,S \rangle} - G_i^{\langle 0,S \rangle}$

- **Marginal productivity metric:** $m_i^S \triangleq \dfrac{f_i^S}{g_i^S}$, provided that $g_i^S \neq 0$

Properties of performance metrics

- **Optimal value function:** $V_i^*(\lambda) = \max_{S \in 2^{\mathcal{N}}} F_i^S - \lambda G_i^S$

Properties of $V_i^*(\lambda)$ (as a function of λ):

1. Convex (hence continuous)
2. Piecewise linear
3. Nonincreasing

- Recall: $v_i^*(\lambda) \triangleq V_i^{\langle 1,* \rangle}(\lambda) - V_i^{\langle 0,* \rangle}(\lambda)$

Properties of $v_i^*(\lambda)$ (as a function of λ):

1. Difference of convex nonincreasing functions (hence continuous)
2. Piecewise linear
3. Need not be monotonic

Motivation of PCL-indexability conditions

- Suppose project is indexable w/ Whittle index satisfying

$$\lambda_1^* < \lambda_2^* < \cdots < \lambda_n^*$$

- Write $S_z \triangleq \{j \in \mathcal{N} : j > z\}$ (active set of z-policy)
- Note: $S_0 = \mathcal{N}$, $S_i = \{i+1, \ldots, n\}$ for $0 \leqslant i < n$, $S_n = \emptyset$
- Then

$$V_i^*(\lambda) = \begin{cases} F_i^{S_0} - \lambda G_i^{S_0}, & \text{if } \lambda \leqslant \lambda_1^* \\ F_i^{S_1} - \lambda G_i^{S_1}, & \text{if } \lambda_1^* \leqslant \lambda \leqslant \lambda_2^* \\ \quad\vdots & \quad\vdots \\ F_i^{S_{n-1}} - \lambda G_i^{S_{n-1}}, & \text{if } \lambda_{n-1}^* \leqslant \lambda \leqslant \lambda_n^* \\ F_i^{S_n} - \lambda G_i^{S_n}, & \text{if } \lambda \geqslant \lambda_n^* \end{cases}$$

Motivation of PCL-indexability conditions

- Then

$$V_i^*(\lambda) = \begin{cases} F_i^{S_0} - \lambda G_i^{S_0}, & \text{if } \lambda \leqslant \lambda_1^* \\ F_i^{S_1} - \lambda G_i^{S_1}, & \text{if } \lambda_1^* \leqslant \lambda \leqslant \lambda_2^* \\ \vdots & \vdots \\ F_i^{S_{n-1}} - \lambda G_i^{S_{n-1}}, & \text{if } \lambda_{n-1}^* \leqslant \lambda \leqslant \lambda_n^* \\ F_i^{S_n} - \lambda G_i^{S_n}, & \text{if } \lambda \geqslant \lambda_n^* \end{cases}$$

- By continuity of $V_i^*(\lambda)$, it follows that the equation

$$F_i^{S_{j-1}} - \lambda G_i^{S_{j-1}} = F_i^{S_j} - \lambda G_i^{S_j}$$

i.e.,

$$F_i^{S_{j-1}} - F_i^{S_j} = \lambda \big(G_i^{S_{j-1}} - G_i^{S_j} \big)$$

has a unique root given by $\lambda = \lambda_j^*$

Some relations between performance metrics

For $j \in S^c$,

$$F_i^{S\cup\{j\}} - F_i^S = f_j^S x_{ij}^{1,S\cup\{j\}} = f_j^{S\cup\{j\}} x_{ij}^{0,S}$$

$$G_i^{S\cup\{j\}} - G_i^S = g_j^S x_{ij}^{1,S\cup\{j\}} = g_j^{S\cup\{j\}} x_{ij}^{0,S}$$

where $$x_{ij}^{a,\pi} \triangleq \mathbb{E}_i^\pi\left[\sum_{t=0}^{\infty} 1_{\{A(t)=a\}}\beta^t\right]$$

Equivalently: for $j \in S$,

$$F_i^S - F_i^{S\setminus\{j\}} = f_j^{S\setminus\{j\}} x_{ij}^{1,S} = f_j^S x_{ij}^{0,S\setminus\{j\}}$$

$$G_i^S - G_i^{S\setminus\{j\}} = g_j^{S\setminus\{j\}} x_{ij}^{1,S} = g_j^S x_{ij}^{0,S\setminus\{j\}}$$

Some relations between performance metrics

For $j \in S^c$ (since $x_{jj}^{1,S\cup\{j\}}, x_{jj}^{0,S} > 0$),

$$\operatorname{sgn}\left(G_j^{S\cup\{j\}} - G_j^S\right) = \operatorname{sgn} g_j^S = \operatorname{sgn} g_j^{S\cup\{j\}}$$

Equivalently: for $j \in S$,

$$\operatorname{sgn}\left(G_j^S - G_j^{S\setminus\{j\}}\right) = \operatorname{sgn} g_j^{S\setminus\{j\}} = \operatorname{sgn} g_j^S$$

Some relations between performance metrics

For $j \in S^c$, if $g_j^S \neq 0$ (recall $m_j^S \triangleq f_j^S / g_j^S$),

$$F_i^{S \cup \{j\}} - F_i^S = m_j^S \big(G_i^{S \cup \{j\}} - G_i^S\big) = m_j^{S \cup \{j\}} \big(G_i^{S \cup \{j\}} - G_i^S\big)$$

Hence, $\qquad m_j^S = m_j^{S \cup \{j\}}$

Equivalently: for $j \in S$, if $g_j^S \neq 0$,

$$F_i^S - F_i^{S \setminus \{j\}} = m_j^{S \setminus \{j\}} \big(G_i^S - G_i^{S \setminus \{j\}}\big) = m_j^S \big(G_i^S - G_i^{S \setminus \{j\}}\big)$$

Hence, $\qquad m_j^{S \setminus \{j\}} = m_j^S$

Some relations between performance metrics

For $j \in S^c$, if $g_j^S \neq 0$ (recall $m_j^S \triangleq f_j^S / g_j^S$),

$$F_i^{S \cup \{j\}} - F_i^S = m_j^S \big(G_i^{S \cup \{j\}} - G_i^S\big) = m_j^{S \cup \{j\}} \big(G_i^{S \cup \{j\}} - G_i^S\big)$$

Hence, $\qquad m_j^S = m_j^{S \cup \{j\}}$

- Recall: By continuity of $V_i^*(\lambda)$, it follows that

$$F_i^{S_{j-1}} - F_i^{S_j} = \lambda_j^* \big(G_i^{S_{j-1}} - G_i^{S_j}\big)$$

Hence, if $g_j^S \neq 0$, we have

$$\lambda_j^* = m_j^{S_{j-1}} = m_j^{S_j}$$

Some relations between performance metrics

- Suppose indexable w/ $\lambda_1^* < \cdots < \lambda_n^*$

$$V_i^*(\lambda) = \begin{cases} F_i^{S_0} - \lambda G_i^{S_0}, & \text{if } \lambda \leqslant \lambda_1^* \\ F_i^{S_1} - \lambda G_i^{S_1}, & \text{if } \lambda_1^* \leqslant \lambda \leqslant \lambda_2^* \\ \vdots & \vdots \\ F_i^{S_{n-1}} - \lambda G_i^{S_{n-1}}, & \text{if } \lambda_{n-1}^* \leqslant \lambda \leqslant \lambda_n^* \\ F_i^{S_n} - \lambda G_i^{S_n}, & \text{if } \lambda \geqslant \lambda_n^* \end{cases}$$

- For $\lambda < \lambda_1^*$, $F_1^{S_0} - F_1^{S_1} > \lambda\big(G_1^{S_0} - G_1^{S_1}\big)$
- For $\lambda_1^* < \lambda < \lambda_2^*$, $F_1^{S_0} - F_1^{S_1} < \lambda\big(G_1^{S_0} - G_1^{S_1}\big)$

This implies $G_1^{S_0} > G_1^{S_1}$, i.e., $g_1^{S_0}, g_1^{S_1} > 0$, and $\lambda_1^* = m_1^{S_0}$

Some relations between performance metrics

- Suppose indexable w/ $\lambda_1^* < \cdots < \lambda_n^*$

$$V_i^*(\lambda) = \begin{cases} F_i^{S_0} - \lambda G_i^{S_0}, & \text{if } \lambda \leqslant \lambda_1^* \\ F_i^{S_1} - \lambda G_i^{S_1}, & \text{if } \lambda_1^* \leqslant \lambda \leqslant \lambda_2^* \\ \vdots & \vdots \\ F_i^{S_{n-1}} - \lambda G_i^{S_{n-1}}, & \text{if } \lambda_{n-1}^* \leqslant \lambda \leqslant \lambda_n^* \\ F_i^{S_n} - \lambda G_i^{S_n}, & \text{if } \lambda \geqslant \lambda_n^* \end{cases}$$

- For $\lambda_1^* < \lambda < \lambda_2^*$, $F_2^{S_1} - F_2^{S_2} < \lambda\big(G_2^{S_1} - G_2^{S_2}\big)$
- For $\lambda_2^* < \lambda < \lambda_3^*$, $F_2^{S_1} - F_2^{S_2} < \lambda\big(G_2^{S_1} - G_2^{S_2}\big)$

This implies $G_2^{S_1} > G_2^{S_2}$, i.e., $g_2^{S_1}, g_2^{S_2} > 0$, and $\lambda_2^* = m_2^{S_1} = m_2^{S_2}$

Some relations between performance metrics

- Suppose indexable w/ $\lambda_1^* < \cdots < \lambda_n^*$

$$V_i^*(\lambda) = \begin{cases} F_i^{S_0} - \lambda G_i^{S_0}, & \text{if } \lambda \leqslant \lambda_1^* \\ F_i^{S_1} - \lambda G_i^{S_1}, & \text{if } \lambda_1^* \leqslant \lambda \leqslant \lambda_2^* \\ \quad \vdots & \quad \vdots \\ F_i^{S_{n-1}} - \lambda G_i^{S_{n-1}}, & \text{if } \lambda_{n-1}^* \leqslant \lambda \leqslant \lambda_n^* \\ F_i^{S_n} - \lambda G_i^{S_n}, & \text{if } \lambda \geqslant \lambda_n^* \end{cases}$$

This implies $G_i^{S_{i-1}} > G_i^{S_i}$, i.e., $g_i^{S_{i-1}}, g_i^{S_i} > 0$, and $\lambda_i^* = m_i^{S_{i-1}} = m_i^{S_i}$

Reformulating marginal value function $v_i^*(\lambda)$

- Suppose indexable w/ $\lambda_1^* < \cdots < \lambda_n^*$

$$v_i^*(\lambda) = \begin{cases} F_i^{\langle 1,S_0\rangle} - F_i^{\langle 0,S_0\rangle} - \lambda\big(G_i^{\langle 1,S_0\rangle} - G_i^{\langle 0,S_0\rangle}\big), & \lambda \leqslant \lambda_1^* \\ F_i^{\langle 1,S_1\rangle} - F_i^{\langle 0,S_1\rangle} - \lambda\big(G_i^{\langle 1,S_1\rangle} - G_i^{\langle 0,S_1\rangle}\big), & \lambda_1^* \leqslant \lambda \leqslant \lambda_2^* \\ \quad\vdots & \quad\vdots \\ F_i^{\langle 1,S_{n-1}\rangle} - F_i^{\langle 0,S_{n-1}\rangle} - \lambda\big(G_i^{\langle 1,S_{n-1}\rangle} - G_i^{\langle 0,S_{n-1}\rangle}\big), & \lambda_{n-1}^* \leqslant \lambda \leqslant \lambda_n^* \\ F_i^{\langle 1,S_n\rangle} - F_i^{\langle 0,S_n\rangle} - \lambda\big(G_i^{\langle 1,S_n\rangle} - G_i^{\langle 0,S_n\rangle}\big), & \lambda \geqslant \lambda_n^* \end{cases}$$

i.e.,

$$v_i^*(\lambda) = \begin{cases} f_i^{S_0} - \lambda g_i^{S_0}, & \lambda \leqslant \lambda_1^* \\ f_i^{S_1} - \lambda g_i^{S_1}, & \lambda_1^* \leqslant \lambda \leqslant \lambda_2^* \\ \quad\vdots & \quad\vdots \\ f_i^{S_{n-1}} - \lambda g_i^{S_{n-1}}, & \lambda_{n-1}^* \leqslant \lambda \leqslant \lambda_n^* \\ f_i^{S_n} - \lambda g_i^{S_n}, & \lambda \geqslant \lambda_n^* \end{cases}$$

Reformulation of of marginal value function

- Suppose indexable w/ $\lambda_1^* < \cdots < \lambda_n^*$

$$v_i^*(\lambda) = \begin{cases} f_i^{S_0} - \lambda g_i^{S_0}, & \lambda \leqslant \lambda_1^* \\ f_i^{S_1} - \lambda g_i^{S_1}, & \lambda_1^* \leqslant \lambda \leqslant \lambda_2^* \\ \vdots & \vdots \\ f_i^{S_{n-1}} - \lambda g_i^{S_{n-1}}, & \lambda_{n-1}^* \leqslant \lambda \leqslant \lambda_n^* \\ f_i^{S_n} - \lambda g_i^{S_n}, & \lambda \geqslant \lambda_n^* \end{cases}$$

- Note that, e.g.:

$$v_1^*(\lambda) = f_1^{S_0} - \lambda g_1^{S_0} > 0, \quad \lambda < \lambda_1^*$$

$$v_1^*(\lambda) = f_1^{S_1} - \lambda g_1^{S_1} < 0, \quad \lambda_1^* < \lambda < \lambda_2^*$$

Some implications of submodularity-based conditions

- Suppose indexable w/ $\lambda_1^* < \cdots < \lambda_n^*$

$$v_i^*(\lambda) = \begin{cases} f_i^{S_0} - \lambda g_i^{S_0}, & \lambda \leqslant \lambda_1^* \\ f_i^{S_1} - \lambda g_i^{S_1}, & \lambda_1^* \leqslant \lambda \leqslant \lambda_2^* \\ \vdots & \vdots \\ f_i^{S_{n-1}} - \lambda g_i^{S_{n-1}}, & \lambda_{n-1}^* \leqslant \lambda \leqslant \lambda_n^* \\ f_i^{S_n} - \lambda g_i^{S_n}, & \lambda \geqslant \lambda_n^* \end{cases}$$

$$v_1^*(\lambda) < v_2^*(\lambda) < \cdots < v_n^*(\lambda), \quad \lambda \leqslant \lambda_1^*$$

i.e.,

$$f_1^{S_0} - \lambda g_1^{S_0} < f_2^{S_0} - \lambda g_2^{S_0} < \cdots < f_n^{S_0} - \lambda g_n^{S_0}, \quad \lambda \leqslant \lambda_1^*$$

- This implies the following (not required by PCLI conditions!):

$$(0 <)\ g_1^{S_0} < g_2^{S_0} < \cdots < g_n^{S_0}$$

Further implications of submodularity-based cond.

- Suppose indexable w/ $\lambda_1^* < \cdots < \lambda_n^*$

$$v_i^*(\lambda) = \begin{cases} f_i^{S_0} - \lambda g_i^{S_0}, & \lambda \leqslant \lambda_1^* \\ f_i^{S_1} - \lambda g_i^{S_1}, & \lambda_1^* \leqslant \lambda \leqslant \lambda_2^* \\ \vdots & \vdots \\ f_i^{S_{n-1}} - \lambda g_i^{S_{n-1}}, & \lambda_{n-1}^* \leqslant \lambda \leqslant \lambda_n^* \\ f_i^{S_n} - \lambda g_i^{S_n}, & \lambda \geqslant \lambda_n^* \end{cases}$$

$$v_1^*(\lambda) < v_2^*(\lambda) < \cdots < v_n^*(\lambda), \quad \lambda \geqslant \lambda_n^*$$

i.e.,

$$f_1^{S_n} - \lambda g_1^{S_n} < f_2^{S_n} - \lambda g_2^{S_n} < \cdots < f_n^{S_n} - \lambda g_n^{S_n}, \quad \lambda \geqslant \lambda_n^*$$

- This implies the following (not required by PCLI conditions!):

$$g_1^{S_n} > g_2^{S_n} > \cdots > g_n^{S_n} \ (> 0)$$

PCL-indexability conditions wrt state ordering $1, \ldots, n$

- Wanted: indexability consistently w/ optim. of threshold policies $S_0 = \{1, 2, \ldots\}$, $S_1 = \{2, 3, \ldots\}$, ... $S_n = \emptyset$
- Active set family: $\mathcal{F} \triangleq \{S_0, S_1, \ldots, S_n\}$

PCL($\mathcal{F}$)-indexability conditions:

- (PCLI1) $g_i^S > 0$ for every $i \in \mathcal{N}$, $S \in \mathcal{F}$
- (PCLI2) For MP index $m_i^* \triangleq m_i^{S_{i-1}} = m_i^{S_i}$: $m_1^* \leqslant \cdots \leqslant m_n^*$

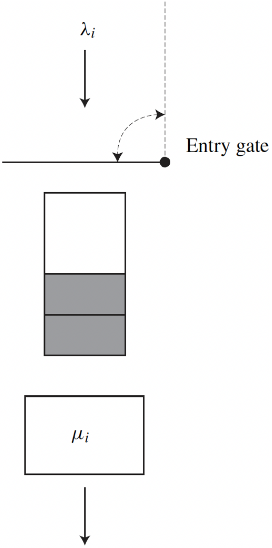
Verification theorem (part (b) more recent)

(a) (PCLI1)+(PCLI2) $\Longrightarrow$ indexable w/ $\lambda_i^* = m_i^*$

(b) Under (PCLI1), indexable w/ $\lambda_i^* = m_i^* \Longleftrightarrow$ (PCLI2)

Example: optimal admission control to a queue

- Holding cost h_i, rejection cost ν, active action = reject



Ex: optimal admission control to a queue (NM '02)

- Write $d_i \triangleq \mu_i - \lambda_i$
- Then, d_i concave nondecreasing $\Longrightarrow$ (PCLI1)
- If, further, h_i is convex nondecreasing $\Longrightarrow$ (PCLI2)
- Such conditions ensure indexability wrt threshold policies
- Sharpest conditions (as far as I know)

PCL-indexability conditions wrt state ordering $1, 2, \ldots$

- Wanted: indexability consistently w/ optim. of threshold policies $S_0 = \{1, 2, \ldots\}$, $S_1 = \{2, 3, \ldots\}$, ... (countable)
- Active set family: $\mathcal{F} \triangleq \{S_0, S_1, \ldots\}$

PCL($\mathcal{F}$)-indexability conditions:

- (PCLI1) $g_i^S > 0$ for every $i \in \mathcal{N}$, $S \in \mathcal{F}$
- (PCLI2) For MP index $m_i^* \triangleq m_i^{S_{i-1}} = m_i^{S_i}$: $m_1^* \leqslant m_2^* \leqslant \cdots$

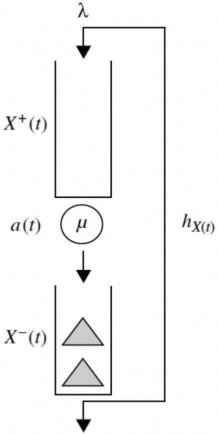
Verification theorem: (part (b) more recent)

(a) (PCLI1)+(PCLI2) $\Longrightarrow$ indexable w/ $\lambda_i^* = m_i^*$

(b) Under (PCLI1), indexable w/ $\lambda_i^* = m_i^* \Longleftrightarrow$ (PCLI2)

Optimal control of a MTO/MTS M/G/1 queue, NM '06

- Net backorder cost h_i, service cost ν, active action = serve
- (PCLI1) holds
- if h_i convex, (PCLI2) holds. Hence, indexable wrt threshold policies



PCL-indexability conditions w/ unknown state ordering

- Typical situation in multidimensional state models

- Need to postulate a structured family of policies, w/ active sets $\mathcal{F}$, which one thinks might be optimal (based on insight)

- Wanted: indexability consistently w/ optimality of $\mathcal{F}$-policies

- Note: need $\emptyset, \mathcal{N} \in \mathcal{F}$, and natural connected properties of $\mathcal{F}$

- (PCLI1) $g_i^S > 0$ for every $i \in \mathcal{N}$, $S \in \mathcal{F}$

- How to define (PCLI2)? Don't know a priori the "right" state ordering

- Will construct it adaptively

Adaptive-greedy($\mathcal{F}$) algorithm and (PCLI2)

- Start w/ $S_0 \triangleq \mathcal{N}$ (which must be in $\mathcal{F}$)
- Pick $i_1 \in \underset{i \in S_0 \colon S_0 \setminus \{i\} \in \mathcal{F}}{\arg\max} m_i^{S_0}; \quad m_{i_1}^* := m_{i_1}^{S_0}; \quad S_1 := S_0 \setminus \{i_1\}$
- Pick $i_2 \in \underset{i \in S_1 \colon S_1 \setminus \{i\} \in \mathcal{F}}{\arg\max} m_i^{S_1}; \quad m_{i_2}^* := m_{i_2}^{S_1}; \quad S_2 := S_1 \setminus \{i_2\}$
- And so on
- $i_n \in \underset{i \colon S_{n-1} \setminus \{i\} \in \mathcal{F}}{\arg\max} m_i^{S_{n-1}}; \; m_{i_n}^* := m_{i_n}^{S_{n-1}}; \; S_n := S_{n-1} \setminus \{i_n\} = \emptyset$

(PCLI2): $m_{i_1}^* \leqslant m_{i_2}^* \leqslant \cdots \leqslant m_{i_n}^*$

Verification theorem (NM 2001, 2002): ((b) more recent)

(a) (PCLI1)+(PCLI2) $\Longrightarrow$ indexable w/ $\lambda_i^* = m_i^*$

(b) Under (PCLI1), indexable w/ $\lambda_i^* = m_i^* \Longleftrightarrow$ (PCLI2)

Extension to countably infinite state

- Start w/ $S_0 \triangleq \mathcal{N}$ (which must be in $\mathcal{F}$)
- Pick $i_1 \in \underset{i \in S_0 :\, S_0 \setminus \{i\} \in \mathcal{F}}{\arg\max} m_i^{S_0}; \quad m_{i_1}^* := m_{i_1}^{S_0}; \quad S_1 := S_0 \setminus \{i_1\}$
- Pick $i_2 \in \underset{i \in S_1 :\, S_1 \setminus \{i\} \in \mathcal{F}}{\arg\max} m_i^{S_1}; \quad m_{i_2}^* := m_{i_2}^{S_1}; \quad S_2 := S_1 \setminus \{i_2\}$
- And so on

(PCLI2): $m_{i_1}^* \leqslant m_{i_2}^* \leqslant \cdots$, w/ $\{i_k : k = 1, 2, \ldots\} = \mathcal{N}$

Verification theorem (NM 2006): ((b) more recent)

(a) (PCLI1)+(PCLI2) $\Longrightarrow$ indexable w/ $\lambda_i^* = m_i^*$

(b) Under (PCLI1), indexable w/ $\lambda_i^* = m_i^* \iff$ (PCLI2)

Further examples of PCL-indexable models

- Finite-buffer delay-/loss-sensitive $M/M/1$ queue (NM ’06)
- Bandits w/ switching costs (NM ’08)
- Finite-horizon bandits (NM ’11)
- Web crawling model (NM ’14)

…

- Age-of-Information scheduling model (NM ’23)

What if the model is not PCL-indexable?

- Example: bandits w/ switching delays (NM '07, '21)

- In NM '07: extension of adaptive-greedy index algorithm which can handle that (relaxes some PCL-indexability requirements, but still ensures indexability)

- It works for bandits w/ switching delays

Fast index computation

- Fast block-implementations of index algorithms (both w/ and w/out PCLs): NM '07, NM '20) for given $\mathcal{F}$

- $O(n^3)$ time if $\mathcal{F} = 2^{\mathcal{N}}$ w/ $\mathcal{N} = \{1, \ldots, n\}$

- But faster if $\mathcal{F} \subset 2^{\mathcal{N}}$

- Recent faster implementation for $\mathcal{F} = 2^{\mathcal{N}}$: Gast, Gaujal and Khun '23

- See also $O(n^3)$ algorithm of Akbarzadeh & Mahajan '22 (for $\mathcal{F} = 2^{\mathcal{N}}$)

A (too) brief history of conservation laws (CLs)

- CLs: fundamental invariance relations on performance metrics for stochastic scheduling models, explain optimality of index policies
- Kleinrock's '65 work CL, multiclass M/G/1 queue: $\sum_{j \in \mathcal{N}} \rho_j \overline{W}_j^\pi \equiv b^{\mathcal{N}}$
- Coffman & Mitrani '80, Gelenbe & Mitrani '80: polyhedral characterization of waiting time performance in multiclass M/G/1 queue

$$\sum_{j \in S} \rho_j \overline{W}_j^\pi \geqslant b^S, \quad S \subset \mathcal{N}$$

- Shanthikumar & Yao '92: Framework of strong CLs
- Bertsimas & NM '96: generalized CLs, (nonrestless) MABP:

$$\sum_{j \in S} g_j^S x_j^\pi \geqslant b^S, \quad S \subset \mathcal{N}; \quad \sum_{j \in \mathcal{N}} g_j^{\mathcal{N}} x_j^\pi \equiv b^{\mathcal{N}}$$

...

Partial CLs (PCLs)

- Bertsimas & NM '96: generalized CLs, (nonrestless) MABP:

$$\sum_{j \in S} g_j^S x_j^\pi \geqslant b^S, \quad S \subset \mathcal{N}; \quad \sum_{j \in \mathcal{N}} g_j^{\mathcal{N}} x_j^\pi \equiv b^{\mathcal{N}}$$

- NM '01, '02: partial CLs, single restless project:

$$G_i^\pi + \sum_{j \in S} g_j^S x_{ij}^{0,\pi} \geqslant G_i^S, \quad \mathcal{N} \neq S \in \mathcal{F}; \quad G_i^\pi + \sum_{j \in \mathcal{N}} g_j^{\mathcal{N}} x_{ij}^{0,\pi} \equiv G_i^{\mathcal{N}}, \ldots$$

- $\mathcal{F} \subseteq 2^{\mathcal{N}}$, but typically one takes $\mathcal{F} \subset 2^{\mathcal{N}}$, i.e., $\mathcal{F}$ is a partial collection of subsets of $\mathcal{N}$

PCLs & indices for multi-gear restless bandits (NM ’22)

- Weber ’07 sketched extension of Whittle index to multi-action bandits
- NM ’08: outlined extension of PCLs for multi-act; NM ’22: full analysis
- A multi-gear project can be operated in multiple gears $a = 0, 1, \ldots, A$
- Higher gears entail larger resource consumption:

$$0 \leqslant q_i^0 < q_i^1 < \cdots < q_i^A, \quad i \in \mathcal{N} = \{1, \ldots, N\}$$

- λ-price problem:

$$\underset{\pi \in \Pi}{\text{minimize}} \ \mathbb{E}_i^\pi \left[\sum_{t=0}^{\infty} \big(h_{s(t)}^{a(t)} + \lambda q_{s(t)}^{a(t)} \big) \beta^t \right]$$

Indexability of multi-gear restless bandits

- λ-price problem:

$$\underset{\pi\in\Pi}{\text{minimize}}\ \mathbb{E}_i^\pi\left[\sum_{t=0}^{\infty}\left(h_{s(t)}^{a(t)}+\lambda q_{s(t)}^{a(t)}\right)\beta^t\right]$$

Definition We call the above multi-gear bandit model indexable if there exist critical resource prices $\lambda_i^{*,a}$ for every state i and active action (gear) $a \geqslant 1$ satisfying $\lambda_i^{*,A} \leqslant \cdots \leqslant \lambda_i^{*,1}$, such that, for any such state and resource price $\lambda \in \mathbb{R}$: (i) action 0 is λ-optimal in state i iff $\lambda \geqslant \lambda_i^{*,1}$; (ii) action $1 \leqslant a \leqslant A-1$ is λ-optimal in state i iff $\lambda_i^{*,a+1} \leqslant \lambda \leqslant \lambda_i^{*,a}$; and (iii) action A is λ-optimal in state i iff $\lambda \leqslant \lambda_i^{*,A}$. We call $\lambda_i^{*,a}$ the model's dynamic allocation index (DAI), viewed as a function of (i,a).

PCL-indexability conditions for multi-gear bandits

Definition We call a multi-gear bandit model PCL-indexable with respect to $\mathcal{F}$-policies, or PCL($\mathcal{F}$)-indexable, if:

- (PCLI1) $g_j^{a-1,a}(S) > 0$ for every policy $S \in \mathcal{F}$, active action $a \geqslant 1$, and state $j \in \mathcal{N}$;
- (PCLI2) Downshift adaptive-greedy algorithm $\mathrm{DS}(\mathcal{F})$ computes the MP index $m_{j_k}^{*,a_k}$ in $\nearrow$ order:

$$m_{j_1}^{*,a_1} \leqslant m_{j_2}^{*,a_2} \leqslant \cdots \leqslant m_{j_K}^{*,a_K}.$$

Theorem

If a multi-gear bandit model is PCL($\mathcal{F}$)-indexable, then it is $\mathcal{F}$-indexable with its DAI being given by its MPI, i.e., $\lambda_j^{*,a} = m_j^{*,a}$.

PCLs for real-state restless bandits

• Real-state restless bandits: sensor scheduling & target tracking POMDP models

• Real-state: probability of channel on (in cognitive radio), posterior variance (target tracking)

• Early results on indexability of real-state restless bandits: Liu & Zhao '08, '10, Le Ny et al. (2008), cognitive radio

• For target tracking, La Scala & Moran '06, Kalman filter model, yet no tools for analysis

• NM '08: outline of PCLs for real-state restless bandits, experiments

• Rigorous PCL framework: NM '15 (arXiv, submitted), '20 (published, *Math. Oper. Res.*)

PCLs for real-state restless bandits

- Notation: $g(x, z) = g_x^{(z,\infty)}$, etc.

(PCLI1) $g(x, z) > 0$ for every state x and threshold z
(PCLI2) MP index $m(x) \triangleq f(x, x)/g(x, x)$ continuous & $\nearrow$
(PCLI3) $F(x, z_2) - F(x, z_1) = \int_{(z_1,z_2]} m(z)\, G(x, dz)$ (Lebesgue–Stieltjes), i.e., $m(\cdot)$: Radon–Nikodym deriv. of $F(x, \cdot)$ wrt $G(x, \cdot)$

Verification theorem (NM '15, '20): (PCLI1)+(PCLI2)+(PCLI3) $\Longrightarrow$ indexable w/ Whittle index $\lambda^*(x) = m(x)$

- Best application to date: Dance & Silander '19 (Kalman filter RBs)
- Empirical application on model extension (submitted '23): (joint work w/ Hao et al.)

Future challenges

• Proving indexability of multi-dimensional discrete-state RB models

• PCL-indexability conditions for multi-dimensional continuous-state RB models

• Implementing & testing adaptive-greedy algorithm for multi-gear bandits

• PCL-indexability conditions for multi-gear continuous-state RB models

• ...

• Anybody wants to join in?

• Note: some references in final slide

Some references

• NM 2001: Restless bandits, partial conservation laws and indexability. *Adv. Appl. Probab*

• NM 2002. Dynamic allocation indices for restless projects and queueing admission control: A polyhedral approach. *Math. Program.*

• NM 2006. Restless bandit marginal productivity indices, diminishing returns and optimal control of make-to-order/make-to-stock M/G/1 queues. *Math. Oper. Res.*

• NM 2007. Dynamic priority allocation via restless bandit marginal productivity indices (with discussion). *TOP*

• NM 2020. A verification theorem for threshold-indexability of real-state discounted restless bandits. *Math. Oper. Res.*

• NM 2022. Multi-gear bandits, partial conservation laws, and indexability. *Mathematics*

• NM 2023. Markovian restless bandits and index policies: A review. *Mathematics*

• NM 2023. Age of information cost minimization with no buffers, random arrivals and unreliable channels: A PCL-indexability analysis. *Mathematics*