Tracking a Stochastic Sequence

Parisa Mansourifard, Bhaskar Krishnamachari, in Collaboration with Tara Javidi

---

### Introduction

- We want to track a stochastic sequence, \( B(t) \)
- The states \( i = 1, \ldots, M \)
- Transition matrix, \( P \)
- Partially Observable Markov Decision Process
- Finite horizon \( T \), time steps \( t = 1, \ldots, T \)
- Our belief vector at time \( t \),
  \[
  b_t = [b_t(1), \ldots, b_t(M)]
  \]
  \[
  b_t(i) = \Pr(B(t) = i), i = 1, \ldots, M
  \]
- Choose an action, \( a \), if higher than \( B(t) \), we will find out the exact amount of \( B(t) \). Otherwise, we can make our belief narrower.
- Evolution of belief vector:
  \[
  b_{t+1} = \begin{cases} T_a b_t P, & \text{if } a \leq B(t) \\ I_{B(t)} P, & \text{if } a > B(t) \end{cases}
  \]
- Current reward:
  \[
  R(B(t); a) = \min(a, B(t)) - C(a - B(t))^+
  \]

### Dynamic Programming

\[
V_t(b_t; a) = \max_{a=1, \ldots, M} V_t(b_t; a), \; \forall t = 1, \ldots, T
\]
\[
V_T(b_T; a) = \bar{R}(b_T; a),
\]
\[
V_t(b_t; a) = \bar{R}(b_t; a) + \beta V_t^f(b_t; a)
\]
\[
V_t^f(b_t; a) = E[V_{t+1}(b_{t+1}) | a]
\]
\[
= \sum_{i=a}^M b_t(i) V_{t+1}(T_a b_t P) + \sum_{i=1}^{a-1} b_t(i) V_{t+1}(I_i P)
\]

### Properties of Value Function

- Convexity with respect to belief vector
  \[
  V_t(\lambda a_1 + (1-\lambda) a_2) \leq \lambda V_t(a_1) + (1-\lambda) V_t(a_2)
  \]
  \( 0 \leq \lambda \leq 1 \)
- Monotonically increase of future expected reward
  \[
  V_t^f(b_t; a_1) \geq V_t^f(b_t; a_2), \; a_1 \geq a_2
  \]

### Myopic Policy

ignoring the impact of the current action on the future reward, myopic policy is given by
\[
a_t^{\text{Myopic}}(b_t) = \arg \max_{a=1, \ldots, M} \bar{R}(b_t; a)
\]
\[
= \min\{a = 1, \ldots, M | \sum_{i=1}^a b_t(i) \geq \frac{1}{1+C}\}
\]

### Optimal Policy

Our goal is to find the optimal policy or prove that it has a threshold structure.
\[
a_t^{\text{Optimal}} = \arg \max_{a=1, \ldots, M} V_t(b_t; a)
\]
Bounds on Optimal Action:
\[
a_t^{\text{Myopic}} \leq a_t^{\text{Optimal}} \leq \max\{i = 1, \ldots, M | b_t(i) \neq 0\}
\]
Future work: Find a tighter upper bound

---

parisama@usc.edu