Optimal stopping policies for weed removal in precision farming

Assignment 2, 4SC000, TU/e, 2022-2023

In the context of precision agriculture, consider a set of N binary states $x_{t,i}$, $i \in \{1, 2, \ldots, N\}$ each representing if there is weed ($x_{t,i} = 1$) or not ($x_{t,i} = 0$) at time t in one of N subareas of a strip of potato crops (see figure below). Weed can simply appear in a given subarea or spread from one of the neighboring subareas, which is summarized by, for $a, b \in \{0, 1\}$,

$$\text{Prob}[x_{t+1,l} = 1 | x_{t,l-1} = a, x_{t,l} = 0, x_{t,l+1} = b] = r_{ab},$$

when $l \notin \{1, N\}$, with $0 < r_{00} < r_{10} = r_{01} < r_{11} < 1$ and $\text{Prob}[x_{t+1,1} = 1 | x_{t,2} = a, x_{t,1} = 0] = \text{Prob}[x_{t+1,N} = 1 | x_{t,N-1} = a, x_{t,N} = 0] = r_a$, $a \in \{0, 1\}$, with $0 < r_0 = r_{00} < r_1 = r_{10} < 1$. When a subarea has weed, it remains until an intervention, $\text{Prob}[x_{t+1,i} = 1 | x_{t,i} = 1] = 1$ for every i. At control or intervention times, labeled by $\sigma_t = 1$, the state is reset to

$$x_{t,1} = 0, \; \ldots \; x_{t,N} = 0, \tag{1}$$

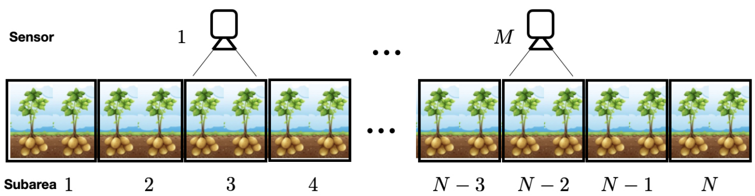
(weed is removed) and cost δ is paid per intervention. Otherwise ($\sigma_t = 0$) weed evolves as described above and there is a cost $g_A(x_{t,1}, \ldots, x_{t,N}) = \sum_{i=1}^{N} g_i(x_{t,i})$, where, for some constant d, representing the cost of having weed between t and $t+1$, $g_i(1) = d$ and $g_i(0) = 0 \quad \forall i \in \{1, \ldots, N\}$. At time zero there is no weed, i.e., $x_{0,i} = 0, \forall i \in \{1, \ldots, N\}$.

Assuming for now that the full state is available, the problem is to find an optimal stationary intervention policy for $\sigma_t = \mu(x_{t,1}, \ldots, x_{t,N})$ to minimize the following average cost

$$J_s = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\left[g_A(x_{t,1}, \ldots, x_{t,N}) + \delta \sigma_t\right].$$

Assume further that at least one intervention must take place in each interval of length $\bar{h}$, i.e., if $\sigma_t = 1$ and $\sigma_{t+1} = 0, \ldots, \sigma_{t+\bar{h}-1} = 0$, then, necessarily, $\sigma_{t+\bar{h}} = 1$. As explained in the appendix, this can be converted into an optimal stopping time problem involving a search over a parameter β. For fixed β such an optimal stopping time problem, which needs to be considered only in the interval $t \in \{0, 1, 2, \ldots, \bar{h} - 1\}$, and, assuming $x_{0,i} = 0, \forall i \in \{1, \ldots, N\}$, is given by

$$J_{\text{stop}} = \min_{\sigma_t = \mu(x_{t,1}, \ldots, x_{t,N})} \mathbb{E}[\delta + \sum_{t=0}^{\tau-1} (g_A(x_{t,1}, \ldots, x_{t,N}) - \beta)]$$



where $\tau = \min\{t \in \{1, \dots, \bar{h}\} | \sigma_t = 1\}$ and β must be found (by line search) such that for the optimal policy the cost of this stopping time problem is zero ($J_{\text{stop}} = 0$). Such a value leads to the minimum original cost $\beta = J_s$ for the optimal policy.

Assignment 2.1 Program a matlab function

```
[Js,mu,avtau] = optimalinterventionpolicy(delta,d,r,hbar,N);
```

that given δ, d, $r = [r_{00} \quad r_{01} \quad r_{11}]$, $\bar{h}$, and N provides the value of $\beta = J_s$ coinciding with the optimal cost J_s that leads to $J_{\text{stop}} = 0$, a matrix mu $= [M_{ij}]_{1 \le i \le 2^N, 1 \le j \le \bar{h}-1}$ specifying the optimal policy, and the average inter-intervention time avtau $= \mathbb{E}[\tau]$, $M_{i,j} = 0$ if $\sigma_t = 0$, $M_{i,j} = 1$ if $\sigma_t = 1$ for the optimal decision when $t = j$ and $i = \text{binx} + 1$ where binx is the decimal value corresponding to the binary value $x_{t,N} x_{t,N-1} \cdots x_{t,1}$; e.g., if $N = 3$, $x_{t,3} = 1$, $x_{t,2} = 1$, $x_{t,3} = 0$ then binx $= 6$.

Consider now that the full state is not available. A single sensor measures only one subarea $\ell \in \{1, \dots, N\}$. It provides $y_t \in \{0, 1\}$ and an error probability is assumed

$$\text{Prob}[y_t = 1 | x_{t,\ell} = 0] = \text{Prob}[y_t = 0 | x_{t,\ell} = 1] = e_p.$$

Let $\mathcal{I}_t = \{y_0, y_1, \dots, y_t\}$ be the sensor information up to time t.

While it is still possible to compute the optimal policy for this case (when N is small, otherwise it becomes computationally intractable), here we restrict ourselves to evaluating the following intuitive threshold policy: if the *expected running* cost (given the sensor information) at a given time t is beyond a given threshold, intervene, i.e.,

$$\sigma_t = \begin{cases} 1 \text{ if } \mathbb{E}[g_A(x_{t,1}, \dots, x_{t,N}) | \mathcal{I}_t] > \gamma \\ 0 \text{ otherwise.} \end{cases}$$

Assignment 2.2 Program a matlab function

```
[Js,avtau] = thresholdinterventionpolicy(delta,d,r,hbar,N,ep,ell,
gamma);
```

that given δ, d, $r = [r_{00} \quad r_{01} \quad r_{11}]$, $\bar{h}$, N, the sensor probability of error e_p, the label ℓ of the statemeasured by the sensor, and γ, provides the values of cost J_s and average inter-intervention time avtau $= \mathbb{E}[\tau]$ for the threshold policy.

Appendix

Considering the control input $u_t = \sigma_t + 1$, we can write the problem in the canonical form

$$x_{t+1} = \underline{f}\left(x_t, u_t, w_t\right)$$

$$J_s = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\left[\underline{g}\left(x_t, u_t\right)\right]$$

where, e.g., $\underline{g}\left(x_t, u_t\right) = \begin{cases} g_A(x_{t,1}, \ldots, x_{t,N}) \text{ if } u_t = 1 \\ \delta \text{ if } u_t = 2 \end{cases}$ and by convention $u_0 = 2$, $s_0 = 0$

Let us partition the cost

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\left[\underline{g}\left(x_t, u_t\right)\right] = \lim_{T\to\infty} \frac{1}{T} \sum_{\ell=0}^{L(T)-1} \mathbb{E}[\sum_{t=s_\ell}^{s_{\ell+1}-1} \underline{g}\left(x_t, u_t\right)] + \lim_{T\to\infty} \frac{1}{T} \sum_{t=s_{L(T)}}^{T-1} \mathbb{E}[\underline{g}\left(x_t, u_t\right)] \quad (2)$$

where $L(T) = \max\{\ell | s_\ell \leq T-1\}$. Since the state resets at each time s_ℓ, we have for every ℓ

$$\mathbb{E}[\sum_{t=s_\ell}^{s_{\ell+1}-1} \underline{g}\left(x_t, u_t\right)] = \mathbb{E}[\sum_{t=s_0}^{s_1-1} \underline{g}\left(x_t, u_t\right)] = \delta + \mathbb{E}[\sum_{t=0}^{\tau-1} g_A(x_{t,1}, \ldots, x_{t,N}))]. \quad (3)$$

Then the cost boils down to

$$\lim_{T\to\infty} \frac{L(T)}{T} \mathbb{E}[\sum_{t=s_0}^{s_1-1} \underline{g}\left(x_t, u_t\right)] = \frac{1}{\mathbb{E}[\tau]} \mathbb{E}[\delta + \sum_{t=0}^{\tau-1} g_A(x_{t,1}, \ldots, x_{t,N})]$$

where $\mathbb{E}[\tau] = \mathbb{E}[s_1]$ and we used the fact that $\lim_{T\to\infty} \frac{L(T)}{T} = \frac{1}{\mathbb{E}[\tau]}$ and the fact that, as $T \to \infty$, the last term in (2) converges to zero.

Suppose now that we can find a policy for σ_t (or equivalently a policy able to determine whether $s_1 = \tau > k$ or $s_1 = \tau = k$ at each time step k) such that

$$\frac{1}{\mathbb{E}[\tau]} \mathbb{E}[\sum_{t=0}^{\tau-1} \underline{g}\left(x_t, u_t\right)] \leq \beta \quad (4)$$

Note that there are finitely many possible policies and as β decreases, the number of policies that satisfy (4) decreases. Now, when β coincides with the optimal cost, only the optimal policy (or optimal policies in case it is not unique) satisfies (4). This (4) can be rewritten (is equivalent) to

$$\mathbb{E}[\delta + \sum_{t=0}^{\tau-1} (g_A(x_{t,1}, \ldots, x_{t,N}) - \beta)] \leq 0.$$

Thus, if β goes below the optimal cost, then no policy satisfies (4) and for all the policies

$$\mathbb{E}[\delta + \sum_{t=0}^{\tau-1} (g_A(x_{t,1}, \ldots, x_{t,N}) - \beta)] > 0.$$

These remarks allow to conclude the following: consider the modified cost

$$\mathbb{E}[\delta + \sum_{t=0}^{\tau-1} (g_A(x_{t,1}, \ldots, x_{t,N}) - \beta)]$$

and iteratively find the optimal policy considering this cost. If the resulting cost is negative, one can still decrease β, if it is positive one must increase β (perform bisection search). When β is such that this auxiliary cost is zero, the optimal policy considering this auxiliary cost is the same as the optimal cost for the problem of minimizing (3), which in turn is also an optimal policy for the original average cost problem.