A Stochastic Assignment Problem

In the paper [1] assume the case of two boxes and define the state of the system as residual quota of the boxes. We want to find the optimal policy by a dynamic programming approach.

Let define $N(S)$ = minimum expected discounted (with factor $\alpha \in (0,1)$) cost in state S. We consider a cost of 1 for each ball arriving. The state for a two box scenario can be like $S = (m_i, m_j, X_i, X_j)$ in which m_i shows the remaining quota of box i and X_i is i'th element of the incoming ball's eligibility vector. Then

$$N(S) = \min_{a \in \mathcal{A}(\mathcal{S})} \{C(S,a) + \alpha \sum_{X'} P(X') N(m_{i'}, m_{j'}, X'_i, X'_j)\}$$

where $\mathcal{A}(\mathcal{S})$ is the set of eligible actions that can be done given state S, note that doing nothing is always a member of this set. $C(S,a)$ represents the cost of taking action a when state is S, which is 1 here, and $m_{i'}$ is the quota of box i after taking action a and X' is the eligibility of the next ball. We can solve this recursive relation using standard dynamic programming in polynomial time in m_i and m_j.

When a ball comes in we solve the previous relation for all the pairs of boxs that the ball is eligible for and the box that gets more wins (gets the ball paired with another box), gets the ball. We want to compare this approach with the policy improvemnt in [1]. Note that the trivial cases are handled easily. For example if a ball is just eligible for one box and that box has remaining quota, then the ball goes into that one and no games are needed. We call the number of balls achieved with our approach as $E[N_{\pi_{\tilde{h}}}]$ and use the same notation as paper [1].

Example: For $p = [0.1,\ 0.15,\ 0.2,\ 0.25,\ 0.3,\ 0.35,\ 0.4,\ 0.45,\ 0.5]$; $m = [5, 8, 10, 12, 15, 18, 20, 23, 26]$, the following compares simulation based estimators of $E[N_{\pi_h}]$, $E[N_{\pi_h *}]$ and $E[N_{\pi_{\tilde{h}}}]$ with the lower bound on the optimal $E[N]$.

$E[N_{\pi_h}]$= 146.1140, $\sigma^2 = 7.9322$
$E[N_{\pi_h *}]$= 144.8000, $\sigma^2 = 12.6222$
$L.B.\{E[N]\}$= 143.5344,

Rep,	Time,	$E[N_{\pi_{\tilde{h}}}]$,	σ^2
10,	10.32,	149.5,	4.385
20,	19.9107,	149.4,	4.422
100,	117.07,	151.1,	1.3135

References

[1] David T. Wu and Sheldon M. Ross. A stochastic assignment problem. *Wiley Online Library*, 2014.