Dynamic Knapsack Problem

KV

February 2021

Contents

1 Questions/Comments

1.1 Questions

- NEED TO MOVE uv TO STATES AND ADJUST ALL THE RELEVANT EQUATIONS FOR IT
- Need to change generating an action section (4.7)

1.2 Comments

-

2 Assumptions

- Assume all new patients are arriving at the beginning of the period
  - Assume arrivals follow poisson distribution
- It is assumed that there is no difference between how long a patient waits for their appointment within some limit (Tl).
  - Let's assume our limit is 10 periods, there is no difference in terms of cost between patients who has been waiting for 1 period and patients waiting 9 period for their appointment (assuming they are the same category).
  - However, if a patient waits for 11 period for their appointment there should be some kind of penalty for the wait
  - The penalty comes from increasing patient complexity. Higher complexities require more resources and are thus more costly overall in the system
- We assume patient complexity transitions follow binary distribution - each period a patient can become more complex with a certain probability
  - While patient is waiting less that Tl, the transition probability is 0.
  - After Tl, transition probability is some arbitrary value
- It is assumed if a patient is scheduled in period one, it means they are served "immediately", regardless if in practice the appointment is at the beginning or end of the period.
- We distinguish between two types of reschedules: good and bad reschedules
  - Good reschedules are reschedules where a patient is rescheduled to an earlier period
  - Bad reschedules are reschedules where a patient is rescheduled to a later period
- We assume that only specific reschedules are allowed (to simplify the model, and remove redundancies)
  - Good reschedules are allowed from any period after 2 into period 1
  - Bad reschedules are only allowed from period 1 to period 2
- It is assumed that there is a certain default expected number of PPE units available for all periods
  - However, in the period 1, there is some random deviation from the expected number of units
  - This random deviation follows some uniform distribution

3 MDP Model

3.1 Decision Epochs

Decisions are made at the beginning of each time period (will be weeks)

3.2 State Space

State space is defined by current units available for various PPEs for future periods, amount of units already used for various PPEs, current patient waitlist, expected period demand, and number of patients already scheduled

$$\vec{S} = (\vec{ue}, \vec{uu}, \vec{uv}, \vec{pw}, \vec{pe}, \vec{ps})$$

- $\vec{ue} = ue_{tp}$ - Expected unist for period t, and PPE p
- $\vec{uu} = uu_{tp}$ - Used units for period t, and PPE p

- $\vec{uv} = uv_{tp}$ - goal variable, violation on number of resources used for period t, of PPE p
- $\vec{pw} = pw_{mdc}$ - Number of patients of complexity d, CPU c, on a wait list for m periods
- $\vec{pe} = pe_{dc}$ - Number of patients of complexity d, CPU c expected to arrive this period
- $\vec{ps} = ps_{tmdc}$ - Number of patients of complexity d, CPU c, scheduled to period t, who have been on the waitlist for m periods (m of 0 stands for people who have just arrived)

3.3 Action Sets

3.3.1 Description

At the beginning of each period, decision maker must reschedule appointments as necessary (if patient complexity increased and too much PPE is being used, or if expected units of PPE have changed negatively). And decision maker must also schedule patients to surgeries

$$\vec{A} = (\vec{sc}, r\vec{sc})$$

- $\vec{sc} = sc_{tmdc}$ - Number of patients of difficulty d, CPU c, who have been in wait list for m periods, to schedule in period t (m of 0 stands for people who have just arrived)
- $r\vec{sc} = rsc_{tt'mdc}$ - Number of patients of difficulty d, CPU c, who have been on the waitlist for m periods, to reschedule from period t to period t'

3.3.2 Action Constraints

- Total number of PPE units cannot be exceeded
  - (U_{pdc} - usage of PPE p per patient difficulty d, CPU c)

$$\sum_{mdc}(sc_{tmdc})U_{pdc} <= (ue_{tp} - uu_{tp}) + uv_{tp} \quad \forall tp \tag{1}$$

- Bounds on when reschedules are allowed

$$rsc_{tt'mdc} = 0 \quad \forall t \in \{2...T\}, t' \in \{2...T\}mdc \tag{2}$$

$$rsc_{tt'mdc} = 0 \quad \forall t \in \{1\}, t' \in \{3...T\}mdc \tag{3}$$

$$rsc_{tt'mdc} = 0 \quad \forall tt'mdc, \text{where } t = t' \tag{4}$$

- Cap on max schedule/reschedule wait time

$$\sum_{tmdc} sc_{tmdc} = 0 \quad \text{where } t + m - 1 > 2Tl \tag{5}$$

$$\sum_{tt'mdc} rsc_{tt'mdc} \quad \text{where } t' + m + 1 > 2Tl \tag{6}$$

- number of people scheduled/rescheduled must be consistent

$$\sum_{t'} rsc_{tt'mdc} \leq ps_{tmdc} \quad \forall tmdc \tag{7}$$

$$\sum_{t} sc_{t0dc} \leq pe_{dc} \quad \forall dc \tag{8}$$

$$\sum_{t} sc_{tmdc} \leq pw_{mdc} \quad \forall m \in \{1...M\}, dc \tag{9}$$

3.4 Transition Probabilities

3.4.1 Uncertainty Sources

1. Number of patients arriving this period - pe_{dc}
   - let's assume pea_{dc} - is the random variable that represents the number of patients arrived this period. It follows a poisson distribution.
2. Transition between patient difficulties within the wait list - pw_{mdc}
   - let's assume pwt_{mdc} is the random variable that represents the number of patients of priority d, CPU c, that have been waiting for m period, that have moved a more complex category. It follows binary distribution.
3. Transition between patient difficulties within the scheduled list - ps_{tmdc}
   - let's assume pst_{tmdc} is the random variable that represents the number of patients of priority d, CPU c, that have been waiting for m period, that have been scheduled into period t, that have moved a more complex category. It follows binary distribution.
4. Amout of expected units of PPE resource for the next time period - ue_{1p}
   - let's assume ued_d is the random variable that represents the deviation of PPE units from the expectation for the next period only. It follows some uniform distribution.
   - let's assume uen_d is the default value to be used for expected number of PPE units p per period

3.4.2 Transition Constraints

1. Transition from $\vec{ue}$ to $\vec{ue'}$ - Expected Units of PPE

$$\begin{aligned} ue'_{1p} =& ue_{2p} + ued_p + ue_{1p} - uu_{1p} - \\ & - \sum_{mdc} sc_{1mdc} U_{pdc} + \sum_{t'mdc} rsc_{1t'mdc} U_{pdc} - \sum_{tmdc} rsc_{t1mdc} U_{pdc} \quad \forall p \end{aligned} \tag{10}$$

$$ue'_{t-1,p} = ue_{tp} \quad \forall t \in \{3...T\}, p \tag{11}$$

$$ue'_{Tp} = uen_p \quad \forall p \tag{12}$$

2. Transition from $\vec{uu}$ to $\vec{uu'}$ - Used Units of PPE

$$\begin{aligned} uu'_{t-1,p} =& uu_{tp} + \sum_{mdc} (sc_{tmdc} U_{pdc}) - \\ & - \sum_{t'mdc} (rsc_{tt'mdc} U_{pdc}) + \sum_{tmdc} (rsc_{tt'mdc} U_{pdc}) + \\ & + \sum_{mdc} pst_{tmdc} (U_{pd+1c} - U_{pdc}) \quad \forall t \in \{2...T\}, p \end{aligned} \tag{13}$$

$$uu'_{Tp} = 0 \quad \forall p \tag{14}$$

3. Transition from $\vec{uv}$ to $\vec{uv'}$ - Violation of PPE units

$$uv'_{tp} \geq uu'_{tp} - ue'_{tp} \quad \forall tp \tag{15}$$

$$uv'_{tp} \geq 0 \quad \forall tp \tag{16}$$

$$uv'_{tp} = \max\{0, uu'_{tp} - ue'_{tp}\} \quad \forall tp \tag{17}$$

4. Transition from $\vec{pe}$ to $\vec{pe}'$ - Expected number of patients for this month

$$pe_{dc} = pea_{dc} \quad \forall dc \tag{18}$$

5. Transition from $\vec{pw}$ to $\vec{pw}'$ - Flow of patients between difficulties/scheduling/cancelling for waitlist

$$pw'_{1dc} = pe_{dc} - \sum_{t} sc_{t0dc} \quad \forall dc \tag{19}$$

$$pw'_{m+1,dc} = pw_{mdc} - \sum_{t} sc_{tmdc} + pwt_{m,d-1,c} - pwt_{mdc} \quad \forall m \in \{1...M-2\}, dc \tag{20}$$

$$pw'_{Mdc} = \sum_{M-1}^{M} pw_{mdc} - \sum_{t,M-1}^{M} sc_{tmdc} + \sum_{M-1}^{M} pwt_{m,d-1,c} - \sum_{M-1}^{M} pwt_{mdc} \quad \forall dc \tag{21}$$

6. Transition from $\vec{ps}$ to $\vec{ps}'$ - Flow of patiensts between difficulties/scheduling/cancelling for scheduled appointments

$$\begin{aligned} ps'_{t-1,m+1,dc} =& ps_{tmdc} + \sum_{t} sc_{tmdc} - \sum_{t'} rsc_{tt'mdc} + \sum_{t} rsc_{tt'mdc} + \\ &+ pst_{tm,d-1,c} - pst_{tmdc} \quad \forall t \in \{2...T\}, m \in \{0...M-2\}, dc \end{aligned} \tag{22}$$

$$\begin{aligned} ps'_{t-1,Mdc} =& \sum_{M-1}^{M} ps_{mdc} - \sum_{t,M-1}^{M} sc_{tmdc} - \sum_{t'M-1}^{M} rsc_{tt'mdc} + \\ & \sum_{t,M-1}^{M} rsc_{tt'mdc} + \sum_{M-1}^{M} pst_{tm,d-1,c} - \sum_{M-1}^{M} pst_{tm,d,c} \quad \forall t \in \{2...T\}, dc \end{aligned} \tag{23}$$

$$ps'_{Tmdc} = 0 \quad \forall mdc \tag{24}$$

3.5 Costs

Cost will come from two source:

- cost of waiting (cw) (comes from 2 things)
- cost of canceling (cc)
- goal variable (to eliminate constraint violation, but still allow it if necessary)

$$\begin{aligned} \vec{C} = c(\vec{pw}, \vec{sc}, \vec{rsc}, \vec{uv}) =& \sum_{mdc} cw_m (pw_{mdc} - \sum_{t} sc_{tmdc}) + \\ &+ cc \sum_{tt`mdc, \text{where } t'>t} rsc_{tt'mdc} - cc \sum_{tt`mdc, \text{where } t'<t} rsc_{tt'mdc} + \\ &+ M(uv_{tp}) \end{aligned} \tag{25}$$

cw_m is computed as follows (val is arbitrary number that describes cost growth):

$$cw_m = val^m \quad \forall m \tag{26}$$

cc is some arbitrary value

4 LP Model

4.1 Full LP

Given a full MDP model, the equivalent LP would look as follows (Note: there is a lot left out. $\vec{C}$ is defined above, all states, actions, and next states are all defined above). Probability distribution is not explicitly defined, but I am not sure if it is necessary

$$\max_{\vec{v}} \sum \alpha(\vec{S})v(\vec{S}) \tag{27}$$

subject to

$$c(p\vec{w}, \vec{sc}, r\vec{sc}, u\vec{v}) + \gamma \sum_{\vec{p}} p(p\vec{ea}, p\vec{wt}, p\vec{st}, u\vec{ed})v(\vec{S'}|\vec{S}, \vec{A}, p\vec{ea}, p\vec{wt}, p\vec{st}, u\vec{ed}) \geq v(\vec{S}) \quad \forall \vec{S}\vec{A} \tag{28}$$

4.2 Full ADP LP

Let's convert it into ADP. We do that by changing $v(\vec{S})$ to an approximation as follows:

$$\begin{aligned} v(u\vec{e}, u\vec{u}, p\vec{w}, \vec{pe}, \vec{ps}) =& \beta^0 + \sum_{tp} \beta^{ue}_{tp} ue_{tp} + \sum_{tp} \beta^{uu}_{tp} uu_{tp} + \sum_{tp} \beta^{uv}_{tp} uv_{tp} + \\ &+ \sum_{mdc} \beta^{pw}_{mdc} pw_{mdc} + \sum_{dc} \beta^{pe}_{dc} pe_{dc} + \sum_{tmdc} \beta^{ps}_{tmdc} ps_{tmdc} \end{aligned} \tag{29}$$

This gives the following LP

$$\max_{\vec{\beta}} \sum \alpha(u\vec{e}, u\vec{u}, p\vec{w}, \vec{pe}, \vec{ps})v(u\vec{e}, u\vec{u}, p\vec{w}, \vec{pe}, \vec{ps}) \tag{30}$$

Subject to:

$$
\begin{aligned}
&c(\vec{pw}, \vec{sc}, \vec{rsc}, \vec{uv}) + \gamma \sum_{\vec{p}} p(\vec{pea}, \vec{pwt}, \vec{pst}, \vec{ued}) * \\
&\Bigg(\beta^0 + \\
&\Big(\sum_p \beta^{ue}_{1p}(ue_{2p} + ued_p + ue_{1p} - uu_{1p} - \sum_{mdc} sc_{1mdc}U_{pdc} + \sum_{t'mdc} rsc_{1t'mdc}U_{pdc} - \sum_{tmdc} rsc_{t1mdc}U_{pdc})\Big) \\
&\Big(\sum_{t=2,p}^{T-1} \beta^{ue}_{tp} ue_{t+1,p} + \sum_p \beta^{ue}_{Tp} uen_d\Big) + \\
&\Big(\sum_p \beta^{uu}_{Tp} * 0\Big) + \\
&\Big(\sum_{tp}^{T-1} \beta^{uu}_{tp}\Big(uu_{t+1,p} + \sum_{mdc}(sc_{t+1,mdc}U_{pdc}) - \sum_{t'mdc}(rsc_{t+1,t'mdc}U_{pdc}) + \sum_{tmdc}(rsc_{tt'+1,mdc}U_{pdc}) + \\
&\qquad \sum_{mdc} pst_{tmdc}(U_{pd+1c} - U_{pdc})\Big)\Big) + \\
&\Big(\sum_{tp} uv'_{tp}\Big) + \\
&\Big(\sum_{dc} \beta^{pe}_{dc} pea_{dc}\Big) + \Big(\sum_{dc} \beta^{pw}_{1dc}(pe_{dc} - \sum_t sc_{t0dc})\Big) + \\
&\Big(\sum_{m=2,dc}^{M-1} \beta^{pw}_{mdc}(pw_{m-1,dc} - \sum_t sc_{t,m-1,dc} + pwt_{m-1,d-1,c} - pwt_{m-1,dc})\Big) + \\
&\Big(\sum_{dc} \beta^{pw}_{Mdc}(\sum_{M-1}^{M} pw_{mdc} - \sum_{M-1,t}^{M} sc_{tmdc} + \sum_{M-1}^{M} pwt_{m,d-1,c} - \sum_{M-1}^{M} pwt_{mdc})\Big) + \\
&\Big(\sum_{mdc} \beta^{ps}_{Tmdc} * 0\Big) + \\
&\Big(\sum_{t=1,m=1,dc}^{T-1,M-1} \beta^{ps}_{tmdc}\Big(ps_{t+1m-1dc} - \sum_t sc_{t+1,m-1,dc} + pst_{t+1,m-1,d-1,c} - pst_{t+1,m-1,dc} - \\
&\qquad - \sum_{t'} rsc_{t+1,t'm-1,dc} + \sum_t rsc_{t,t'+1,m-1,dc}\Big)\Big) + \\
&\Big(\sum_{t=1dc}^{T-1} \beta^{ps}_{tMdc}\Big(\sum_{M-1}^{M} ps_{t+1,mdc} - \sum_{M-1,t}^{M} sc_{t+1,mdc} + \sum_{M-1}^{M} pst_{t+1,m,d-1,c} - \sum_{M-1}^{M} pst_{t+1,mdc} - \\
&\qquad - \sum_{M-1,t'}^{M} rsc_{t+1,t'mdc} + \sum_{M-1,t}^{M} rsc_{t,t'+1,mdc}\Big)\Big) \\
&\Bigg) \geq v(\vec{ue}, \vec{uu}, \vec{pw}, \vec{pe}, \vec{ps}) \quad \forall \vec{S}\vec{A}
\end{aligned}
$$

$$\vec{\beta} \geq 0 \quad \forall \beta$$

4.3 Expectation ADP LP

Steps for converting the mega constraint $v(\vec{S'})$ into expectation and merging it with $v(\vec{S})$:

- β^0 - no adjustments for expectation for $v(\vec{S'})$

$$\beta^0 - \gamma\beta^0 = (1-\gamma)\beta^0$$

- β^{ue}_{tp}

$$E[ued_p] = 0 \quad \therefore \quad \sum_p \beta^{ue}_{1p}(ue_{2p} + ued_p) \rightarrow \sum_p \beta^{ue}_{1p}(ue_{2p}) \rightarrow \sum \beta^{ue}_{tp} uen_p$$

$$\Bigg(\Bigg(\sum_p \beta^{ue}_{1p}(ue_{1p} - \gamma\Big(uen_p + ue_{1p} - uu_{1p} - \sum_{dc} U_{pdc}(\sum_m sc_{mdc} + \sum_{t'm} rsc_{1t'mdc} - \sum_{tm} rsc_{t1mdc})\Big)\Bigg)+ \Bigg(\sum_{t=2p}^{T} \beta^{ue}_{tp}(ue_{tp} - \gamma(uen_p)\Big)\Bigg)$$

- β^{uu}_{tp} - no adjustments for expectation

$$uu_{Tp} = 0 \quad \therefore \quad \beta^{uu}_{Tp} uu_{Tp} = 0$$

$$\sum_{tp}^{T-1} \beta^{uu}_{tp} uu_{tp} - \gamma \sum_{tp}^{T-1} \beta^{uu}_{tp}\Bigg(uu_{t+1p} + \sum_{mdc}(sc_{t+1mdc}U) - \sum_{tmdc}(rsc_{t+1t'mdc}U) + \sum_{t'mdc}(rsc_{tt'+1mdc}U) + \sum_{mdc} ps_{tmdc} pwp_{mdc}(U_{pd+1c} - U_{pdc})\Bigg)$$

$$\Bigg(\sum_{tp}^{T-1} \beta^{uu}_{tp}\Bigg(uu_{tp} - \gamma\Big(uu_{t+1p} + \sum_{mdc} ps_{tmdc} pwp_{mdc}(U_{pd+1c} - U_{pdc}) + U_{pdc}(\sum_{mdc} sc_{t+1mdc} - \sum_{t'mdc} rsc_{t+1t'mdc} + \sum_{tmdc} rsc_{tt'+1mdc})\Big)\Bigg) + \beta^{uu}_{Tp}(uu_{Tp})\Bigg)$$

- β^{uv}_{tp}

$$E[uv_{tp}] = 0$$

$$\sum_{tp} \beta^{uv}_{tp}(uv_{tp} - \gamma(E[uv_{tp}])) = \sum_{tp} \beta^{uv}_{tp} uv_{tp}$$

- β^{pe}_{dc}

$$\sum_{dc} \beta^{pe}_{dc} pe_{dc} - \gamma \sum_{dc} \beta^{pe}_{dc} E[pea_{dc}] = \sum_{dc} \beta^{pe}_{dc}(pe_{dc} - \gamma E[pea_{dc}])$$

- β^{pw}_{mdc} - Let's assume pwp_{mdc} - is the probability of a patient transitioning to a higher complexity

$$E[pwt_{mdc}] = pw_{mdc} pwp_{mdc}$$

$$\Bigg(\sum_{dc} \beta^{pw}_{1dc} (pw_{1dc} - \gamma(pe_{dc} - \sum_{t} sc_{t0dc})) +$$
$$\sum_{m=2,dc}^{M-1} \beta^{pw}_{mdc} (pw_{mdc} - \gamma\big(pw_{m-1dc}(1-pwp) + pw_{m-1d-1c}pwp - \sum_{t} sc_{tm-1dc}\big) +$$
$$\sum_{dc} \beta^{pw}_{Mdc} (pw_{Mdc} - \gamma \sum_{M-1}^{M} \big(pw_{mdc}(1-pwp) + pw_{md-1c}pwp - \sum_{t} sc_{tmdc}\big)\Bigg)$$

- β^{ps}_{mdc}

$$E[pst_{tmdc}] = ps_{tmdc} pwp_{mdc}$$

$$\Bigg(\sum_{t=1m=1dc}^{T-1M-1} \beta^{ps}_{tmdc} \Big(ps_{tmdc} - \gamma \Big(ps_{t+1m-1dc}(1-pwp) + ps_{t+1m-1d-1c}pwp - \sum_{t} sc_{t+1m-1dc}$$
$$- \sum_{t'} rsc_{t+1t'm-1dc} + \sum_{t} rsc_{tt'+1m-1dc} \Big)\Big) +$$
$$\sum_{t=1dc}^{T-1} \beta^{ps}_{tMdc} \Big(ps_{tMdc} - \gamma \sum_{M-1}^{M} \Big(ps_{t+1mdc}(1-pwp) + ps_{t+1md-1c}pwp - \sum_{t} sc_{t+1mdc}$$
$$- \sum_{t'} rsc_{t+1t'mdc} + \sum_{t} rsc_{tt'+1mdc} \Big)\Big) +$$
$$\sum_{mdc} \beta^{ps}_{Tmdc} \Big(ps_{Tmdc} \Big) \Bigg)$$

Let's say $E[V]$ is the addition of all the red parts above. In that case the ADP model converted to expectation would look as follows:

$$\max_{\vec{\beta}} \sum \beta^0 + \sum \beta^{ue} E[ue] + \sum \beta^{uu} E[uu] + \sum \beta^{uv} E[uv] \sum \beta^{pe} E[pe] + \sum \beta^{pw} E[pw] + \sum \beta^{ps} E[ps] \quad (31)$$

Subject to:

$$E[V] \le c(\vec{pw}, \vec{sc}, r\vec{sc}, \vec{uv}) \quad \forall \vec{S}\vec{A} \quad (32)$$

4.4 Dual of ADP LP

Converting to Dual

$$\min_{w} \sum_{\vec{S}\vec{A}} w(\vec{S}\vec{A}) c(\vec{S}\vec{A}) \quad (33)$$

Subject To:

- β^0 constraints

$$\sum_{\vec{S}\vec{A}} \vec{w}(1-\gamma) = 1 \quad (34)$$

- β^{ue} constraints

$$\sum_{\vec{S}\vec{A}} \vec{w}\Big(ue_{1p} - \gamma\Big(uen_p + ue_{1p} - uu_{1p} - \sum_{dc} U_{pdc}(\sum_{m} sc_{mdc} + \sum_{t'm} rsc_{1t'mdc} - \sum_{tm} rsc_{t1mdc})\Big)\Big) \geq E[ue_{1p}] \quad \forall p \tag{35}$$

$$\sum_{\vec{S}\vec{A}} \vec{w}\Big(ue_{tp} - \gamma(uen_p)\Big) \geq E[ue_{tp}] \quad \forall t \in (2...T)p \tag{36}$$

- β^{uu} constraints

$$\sum_{\vec{S}\vec{A}} \vec{w}\Big(uu_{tp} - \gamma\Big(uu_{t+1p} + \sum_{mdc} ps_{tmdc}pwp(U_{pd+1c} - U_{pdc}) + U_{pdc}(\sum_{mdc} sc_{t+1mdc} - \sum_{t'mdc} rsc_{t+1t'mdc} + \sum_{tmdc} rsc_{tt'+1mdc})\Big)\Big) \geq E[uu_{tp}] \quad \forall T \in \{1...T-1\}p \tag{37}$$

$$\sum_{\vec{S}\vec{A}} \vec{w}(uu_{Tp}) \geq E[uu_{Tp}] \quad \forall p \tag{38}$$

- $\beta^{uv} constraints$

$$\sum_{\vec{S}\vec{A}} \vec{w}(uv_{tp}) \geq 0 \quad \forall tp \tag{39}$$

- β^{pe} constraints

$$\sum_{\vec{S}\vec{A}} \vec{w}(pe_{dc} - \gamma E[pea_{dc}]) \geq E[pe_{dc}] \quad \forall dc \tag{40}$$

- β^{pw} constraints

$$\sum_{\vec{S}\vec{A}} w(pw_{1dc} - \gamma(pe_{dc} - \sum_{t} sc_{t0dc})) \geq E[pw_{1dc}] \quad \forall dc \tag{41}$$

$$\sum_{\vec{S}\vec{A}} w\Big(pw_{mdc} - \gamma\big(pw_{m-1dc}(1-pwp) + pw_{m-1d-1c}pwp - \sum_{t} sc_{tm-1dc}\big)\Big) \geq E[pw_{mdc}] \quad \forall m \in \{2...M-1\}dc \tag{42}$$

$$\sum_{\vec{S}\vec{A}} w\Big(pw_{Mdc} - \gamma \sum_{M-1}^{M} \big(pw_{mdc}(1-pwp) + pw_{md-1c}pwp - \sum_{t} sc_{tmdc}\big)\Big) \geq E[pw_{Mdc}] \quad \forall dc \tag{43}$$

- β^{ps} constraints

$$\sum_{\vec{S}\vec{A}} w\Big(ps_{tmdc} - \gamma\Big(ps_{t+1m-1dc}(1-pwp) + ps_{t+1m-1d-1c}pwp - \sum_{t} sc_{t+1m-1dc} - \sum_{t'} rsc_{t+1t'm-1dc} + \sum_{t} rsc_{tt'+1m-1dc}\Big)\Big) \geq E[ps_{tmdc}] \quad \forall t \in \{t...T-1\}m \in \{1...M-1\}dc \tag{44}$$

$$\sum_{\vec{S}\vec{A}} w\Big(ps_{tMdc} - \gamma \sum_{M-1}^{M} \Big(ps_{t+1mdc}(1-pwp) + ps_{t+1md-1c}pwp - \sum_{t} sc_{t+1mdc} - \sum_{t'} rsc_{t+1t'mdc} + \sum_{t} rsc_{tt'+1mdc}\Big)\Big) \geq E[ps_{tmdc}] \quad \forall t \in \{t...T-1\}dc \quad (45)$$

$$\sum_{\vec{S}\vec{A}} \vec{w}(ps_{Tmdc}) \geq E[ps_{Tmdc}] \quad \forall mdc \quad (46)$$

- State-Action constraints

$$w \geq 0 \quad \forall w \quad (47)$$

4.5 Finding most violated constraints of ADP LP

$$\min_{(\vec{ue},\vec{uu},\vec{pw},\vec{pe},\vec{ps})\in S,(\vec{sc},\vec{rsc})\in A} c(\vec{pw}, \vec{sc}, \vec{rsc}, \vec{uv}) - E[V] \quad (48)$$

Subject to: all constraints in section "Action Constraints"

4.6 Algorithm for solving

To get $\vec{\beta}$ values, which will be used to generate an action follow steps below:

1. Perform a monte-carlo simulation (following some arbitrary policy) to get $E[ue], E[uu], E[pe], E[pw]E[pw]$ (this will only give an approximation as arbitrary policy will likely be different from ADP policy) (if you wanted to, you could resolve for $\vec{\beta}$ again, but replacing $E[ue], E[uu], E[pe], ...$ with the ones that would be generated through the ADP policy)
2. Create an initial feasible set of state-action pairs - $\vec{w}$
3. Solve model in section "Dual of ADP LP" where each state-action pairs in $\vec{w}$ corresponds to a variable and parameters for all the constraints
4. Solve model in section "Finding most violated constraints of ADP LP", where duals from problem in step 3 correspond to $\vec{\beta}$ values.
   - If objective function is less than 0, add solution as a single state-action pair to $\vec{w}$ and go to step 3
   - If objective function is greater than 0, continue to next step
5. Duals from problem in step 3 correspond to final $\vec{\beta}$ values

4.7 Generating an Action

Once $\vec{\beta}$ values have been approximated, you may use the model below to generate a recommended action for a specific state.

$$\min_{\vec{A}} c(\vec{S}, \vec{A}) + E[V] \quad (49)$$

Subject to: all constraints in section "Action Constraints"