# DP Decomposition Approaches

- The Network RM DP

$$V(t,\underline{x}) - V(t-1,\underline{x}) = \sum_{i=1}^{n} \lambda_i(t)\left[P_i - \Delta_i V(t-1,\underline{x})\right]^+$$

where $\Delta_i V(t-1,\underline{x}) = V(t-1,\underline{x}) - V(t-1,\underline{x}-\underline{a}_i)$

$$V(t,0) = V(0,x) = 0$$

$$V(t,\underline{x}) = -\infty \quad \forall\ \underline{x} < 0 \ (\text{feasibility cond}^n)$$

- The 'fluid' LP approximation $\left(\Lambda_i = \int_0^T \lambda_i(s)\,ds\right)$

$$y_i^* \equiv \max \sum_{i=1}^{n} P_i y_i \quad \text{s.t.} \quad \sum_{i=1}^{n} \underline{a}_i y_i \le \underline{C}, \quad 0 \le y_i \le \Lambda_i$$

$$z_j^* \equiv \min \sum_{i\in[n]} \Lambda_i \beta_i + \sum_{j\in[m]} c_j z_j \quad \text{s.t.} \quad \sum_{i=1}^{n} \underline{a}_i z_j + \underline{\beta} \ge \underline{P}, \quad z_j \ge 0,\ \beta_i \ge 0$$

---

- Idea - Use tighter upper bounds on $V(t,x)$ which admit easy-to-compute policies
  - In particular - decomposable bounds across ~~&~~ resources

<u>Approach 1</u> - Use $z_j^*$ to relax all capacity constraints except for $j$

$$\overline{V}^j(T, c) = \max \sum_{i \in [n]} y_i \Big[ p_i - \sum_{\substack{k \in [m] \\ k \neq j}} z_k^* a_i(k) \Big] + \sum_{\substack{k \in [m] \\ k \neq j}} z_k^* c_k$$

$$\text{s.t} \quad \sum_{i \in [n]} a_i(j)\, y_i \leq c_j$$

$$0 \leq y_i \leq \Lambda_i \quad \forall i \in [n]$$

— This is now a single-resource problem for $j$, with 'virtual' prices $\Big\{ p_i - \sum_{\substack{k \in [m] \\ k \neq j}} a_i(k)\, z_k^* \Big\}_{i \in [n]}$

$$v_j(t, x_j) = \max_{\sum a_i(j) u_i \in \{0, 1, \dots, x_j\}} \Big\{ \sum_{i \in [n]} \lambda_j(t) \Big\{ \Big[ p_i - \sum_{k \neq j} a_i(k) z_k^* \Big] u_i + v_j(t-1, x_j - u_i a_i(j)) \Big\} \Big\}$$

- <u>Thm</u> - $\forall\, t \in \{1, 2, \dots, T\}, \forall\, \underline{x}$

$$V(t, x) \leq \min_{j \in [m]} \Big\{ v_j(t, x_j) + \sum_{k \neq j} z_k^* x_k \Big\}$$

(Pf Idea - Induction on $t$)

- <u>Heuristic</u> - Use $\sum_{j \in [m]} v_j(t, x_j)$ to approximate $V(t, \underline{x})$

## Approach 2 - DP Decomposition via Lagrangean penalties

- Let $w_{ij}(t) = \mathbb{1}\{\text{Class } i \text{ alloted resource } j \text{ at time } t\}$

  $w_i(t) = \mathbb{1}\{\text{class } i \text{ is admitted at time } t\}$

Now we can rewrite our DP as:

$$V(t, \underline{x}) = \max_{\{w_{ij}(t)\}} \sum_{i \in [n]} \lambda_i(t) \left[ P_i w_i(t) + V\left(t-1, \underline{x} - \sum_j w_{ij}(t) a_i(j) \underline{e}_j\right)\right]$$

$$\text{s.t} \quad a_i(j)\, w_{ij}(t) \leq x_j \quad \forall j \in [m],\ i \in [n]$$

$$\text{(coupling constraint)} \quad w_{ij}(t) = w_i(t) \quad \forall j \in [m],\ i \in [n]$$

$$w_{ij}(t) \in \{0,1\}, \quad w_i(t) \in \{0,1\}$$

- Idea: We can dualize the coupling constraint via Lagrange multipliers $[\alpha_{ij}(t)\lambda_i(t)]$

$$V^{\alpha}(t, \underline{x}) = \max_{\{w_{ij}(t)\}} \sum_{i \in [n]} \lambda_i(t) \left[ \sum_{j \in [m]} \alpha_{ij}(t) w_{ij}(t) + \left[P_i - \sum_{j \in [m]} \alpha_{ij}(t)\right] w_i(t) + V^{\alpha}\left(t-1, \underline{x} - \sum_{j \in [m]} w_{ij}(t) a_i(j) \underline{e}_j\right)\right]$$

$$\text{s.t} \quad \left.\begin{array}{l} \cancel{\sum} a_i(j)\, w_{ij}(t) \leq x_j \quad \forall i, j \\ w_{ij} \in \{0,1\} \end{array}\right] \triangleq w_j \in U_j(x_j) = \{w_j \in \{0,1\}^n \text{ s.t } a_i(j) w_{ij} \leq x_j \ \forall i\}$$

Note: This decomposes across resources!

- Now consider the single-resource DP recursion

$$\vartheta_j^{\alpha}(t, x_j) = \max_{\substack{\omega_j \in \mathcal{U}_j(x_j) \\ (\text{ie, } \{\omega_j \in \{0,1\}^n \text{ s.t. } \omega_{ij} a_i(j) \le x_j \, \forall i\})}} \left\{ \sum_{i \in [n]} \lambda_i(t) \left\{ \alpha_{ij}(t)\, \omega_{ij}(t) + \vartheta_j^{\alpha}(t-1, x_j - \omega_{ij}(t) \cdot a_i(j)) \right\} \right\}$$

Then $$V^{\alpha}(t, \underline{x}) = \sum_{j \in [m]} \vartheta_j^{\alpha}(t, x_j) + \sum_{i \in [n]} \sum_{\tau=1}^{t} \lambda_i(\tau) \Big[ P_i - \sum_{j \in [m]} \alpha_{ij}(\tau) \Big]^+$$

(Can be shown by induction)

- Thus we can compute $V^{\alpha}(t, \underline{x})$ for any given Lagrange multipliers $\alpha$, via $j$ single-resource DPs.

Moreover, we have $\forall \{\alpha_{ij}(t)\}_{i,j,t} \in \mathbb{R}^{Tmn}$

$$\boxed{V(t, x) \le V^{\alpha}(t, x)} \quad \left( \begin{array}{l} \text{induction} \\ + \text{weak duality} \end{array} \right)$$

- How can we choose good $\alpha$? Optimize!
  - $\alpha^* \equiv \arg\min_{\alpha \in \mathbb{R}^{Tmn}} V_1^{\alpha}(c)$

<u>Claim</u> - $V^{\alpha}(1, c)$ is convex in $\alpha$ (see assgnt)

$\therefore$ We can get good $\alpha$ via convex optimization

# LP approach to DP

- Consider generic Bellman eqn (finite $S$, $A(s)$)

$$V_t(s) = \max_{a \in A(s)} \left[ r_t(s,a) + \sum_{j \in S} P(j|s,a)\, V_{t+1}(j) \right]$$

$$V_{T+1}(s) = 0 \quad \forall\, s \in S$$

- This is equivalent to the following LP (Given we start at $t=1$ in $s_1$)

$$Z^* = \min \quad \tilde{V}_1(s_1)$$

$$\text{s.t} \quad \tilde{V}_t(s) \geq r_t(s,a) + \sum_{j \in S} P_t(j|s,a)\, \tilde{V}_{t+1}(j)$$

$$\forall\, t \in \{1, \dots, T\},\ \forall\, s \in S,\ a \in A(s)$$

$$\tilde{V}_T(s) \geq r_T(s,a) \quad \forall\, s \in S,\ a \in A(s)$$

Note -

1) $Z^* = V_1(s_1)$ (Can be shown via induction on $t$)
2) Decision vars $\equiv \{ V_t(s)\ ;\ t \in [T],\ s \in S \}$
3) Any feasible soln $\tilde{V}_t(s)$ is an upper bound on $V_t(s)$
4) $T|S|$ variables, $T \cdot \left( \sum_{s \in S} |A(s)| \right)$ constraints

- Dual of LP -

$$\max \sum_{t=1}^{T} \sum_{s\in S} \sum_{a\in A(s)} r_t(s,a)\, Y_t(s,a)$$

$$\text{s.t} \quad \sum_{a\in A(s)} Y_t(s,a) = \sum_{j\in S} \sum_{a\in A(j)} P_{t-1}(s|j,a)\, Y_{t-1}(j,a) \qquad \forall\, s\in S,\ t\in[T]$$

$$\sum_{a\in A(s)} Y_1(s,a) = \begin{cases} 1 & \text{if } s = s_1 \\ 0 & \text{if } s \neq s_1 \end{cases}$$

$$Y_t(s,a) \geq 0$$

- $Y_t(s,a) \equiv$ 'probability' of landing in $s$ at time $t$, and taking action $a$
- This has $T|S|$ constraints, $T\left(\sum_{s\in S} |A(s)|\right)$ variables

---

- In general, the LPs given above are too high dimensional to solve
- <u>Idea</u> - Approximate $\tilde{V}_t^{\beta}(s) = \sum_{k=1}^{K} \beta\, \phi_{kt}(s)$

where $\phi_{kt}(\cdot)$ are a set of <u>$K$ basis fns</u> $(K \ll S)$

$\Rightarrow$ $TK$ variables, $TK|A|$ constraints (but can use row generation)

# Approx LP for Network RM

- Let $G \triangleq \{ [c_1] \times [c_2] \times \ldots \times [c_m] \}$

  $U(\underline{x}) \triangleq \{ \omega \in \{0,1\}^n \mid \omega_i A_i \leq \underline{x} \}$

Then LP equivalent formulation $\equiv$

$$\min \; \vartheta(T, c)$$

$$\text{s.t} \quad \vartheta(t, x) \geq \sum_{i \in [n]} \lambda_i(t) \left[ p_i \omega_i + \vartheta(t-1, x - \omega_i A_i) \right]$$

$$\forall t, \forall x, \forall \omega \in U(x)$$

- Decision vars - $\{ \vartheta(t, x) \; ; \; t \in [T], \; x \in G \}$

- $\vartheta(0, x) = 0 \quad \forall \; x \in G$

- Use $\{ \Phi_k(\cdot) \}_{k \in [K]}$ as basis to approximate $\vartheta$

$$\Rightarrow \qquad \min \sum_{k=1}^{K} r_{Tk} \, \Phi_k(c)$$

$$\text{s.t} \quad \sum_{k=1}^{K} r_{tk} \, \Phi_k(x) \geq \sum_{i \in [n]} \lambda_i(t) \left[ p_i \omega_i + \sum_{k=1}^{K} r_{t-1,k} \Phi_k(x - \omega_i A_i) \right]$$

$$\forall t \; \forall x \in G, \forall \omega \in U(x)$$