$$\min_v \; c^T v \qquad c \ge 0$$

$$\text{s.t.} \quad v \ge r_1 + \gamma \cdot P_1 v, \qquad v \ge r_2 + \gamma \cdot P_2 v \qquad v \ge v^*$$

$$(I - \gamma \cdot P_1) v \ge r_1$$
$$(I - \gamma P_2) v \ge r_2$$

ALP

Primal

$$\min_w \; c^T \Phi w$$
$$\text{s.t.} \quad \underbrace{\begin{bmatrix} (I-\gamma P_1)\Phi \\ (I-\gamma P_2)\Phi \end{bmatrix}}_{A} w \ge \begin{matrix} r_1 \\ r_2 \end{matrix}$$

- Linear program solved

$$A = \begin{bmatrix} I - \gamma P_1 \\ I - \gamma P_2 \end{bmatrix} \qquad r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$$

$$r_1 = \begin{bmatrix} r(s_1,a_1) \\ r(s_2,a_1) \\ \vdots \end{bmatrix}$$

$$P_1 = \begin{bmatrix} p(s_1,a_1,\cdot)^T \\ p(s_2,a_1,\cdot)^T \\ \vdots \end{bmatrix}$$

$$\min \; c^T \Phi w \qquad \text{s.t.} \quad A\Phi w \ge r$$

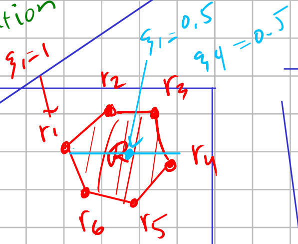
$$\max_u \; r^T u \qquad \text{s.t.} \quad \Phi^T A^T u = \Phi^T c, \quad u \ge 0$$

Reward function r^1, r^2

$$\max_\pi \min_{r \in R} \rho(\pi, r)$$

Robust optimization

$$\max_u \; \min_{r \in R} r^T u$$
$$\text{s.t.} \quad \sum_{a \in A} (I - \gamma P_a^T) u_a = c$$
$$u_a \ge 0$$



$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\min_\xi \; u^T \left(\sum_{i=1}^n \xi_i r_i \right) \qquad \text{s.t.} \quad \sum_i \xi_i = 1, \quad \xi \ge 0$$

$$\max_z \; z \qquad \text{s.t.} \quad z \le r^T u \quad \forall r \in R$$

$$\max_{u,z} \; z$$
$$\text{s.t.} \quad z \le r^T u \quad \forall r \in R$$
$$\sum_{a \in A} (I - \gamma P_a^T) u_a = c$$
$$u_a \ge 0$$
$$R = \{r\}$$

$$\max_{u,z} \; z$$
$$\text{s.t.} \quad z \le r^T u \quad \forall r \in R = \{r_1, r_2, r_3\}$$
$$A^T u = c$$
$$u \ge 0$$

$$R = \begin{bmatrix} | & | & | \\ r_1 & r_2 & r_3 \\ | & | & | \end{bmatrix}$$

dual

$$\min_{w,\xi} \; c^T \Phi w$$
$$\text{s.t.} \quad u: \; A \Phi w \ge R\xi$$
$$z: \; 1^T \xi = 1$$
$$\xi \ge 0$$

$$\Phi = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\max_{u,z} \; z$$
$$z - r^T u \le 0$$
$$\text{s.t.} \quad \xi: \; z \le r^T u \quad r \in R$$
$$w: \; \Phi^T A^T u = \Phi^T c$$
$$u \ge 0$$
$$\xi: \; z \cdot 1 - R^T u \le 0$$

$\circ \quad \max_{\pi} \min_{\substack{p \in \mathcal{P} \\ r \in \mathcal{R}}} \rho(\pi, P, r)$

$$\min_{w, \xi} \; c^T \Phi w$$
$$\text{s.t.} \; A \Phi w \geq \hat{\Phi} \tilde{R} \xi$$
$$1^T \xi = 1$$
$$\xi \geq 0$$

state-action features

reward weights

state features

$\max_{\pi} \min_{p \in \mathcal{P}} \rho(\pi, P, r) \approx$ LP

NP hard

$$\hat{\Phi} = \begin{bmatrix} \Phi & 0 \\ 0 & \Phi \end{bmatrix}$$

$$v(s) = \max_{a \in \mathcal{A}} \left(r_{s,a} + \gamma \, p_{s,a}^T v \right)$$

$$\boxed{v(s) = \max_{a \in \mathcal{A}} \min_{p \in \mathcal{P}_{s,a}} \left(r_{s,a} + \gamma \, p^T v \right)}$$

Can solve it

$\Rightarrow$???

2. Infinite horizon

Fininite horizon

R-LP

$$\max \; z$$
$$\text{s.t.} \; z \leq r^T u, \; r \in \mathcal{R}$$
$$A^T u = c$$
$$u \geq 0$$

$<$

LP

$$\max_u \; r^T u$$
$$\text{s.t.} \; A^T u = c$$
$$u \geq 0$$

$<$

ALP

$$\max_u \; r^T u$$
$$\text{s.t.} \; \Phi^T A^T u = c$$
$$u \geq 0$$

O

ALP :

$$\min \quad c^T \Phi w$$
$$\text{s.t.} \quad A \Phi w \geq r$$

$$L(w, u) = \underline{c^T \Phi w} + \underline{u^T (r - A \Phi w)}$$

$$q(u) = \inf_w \underbrace{\left(c^T \Phi - u^T A \Phi\right)}_{\neq 0} w + u^T r$$

$$\max_{u \geq 0} q(\bar{u}) = \boxed{\begin{aligned} &\max_u \quad u^T r \\ &\text{s.t.} \quad u \geq 0 \\ &\qquad [u^T A \Phi = c^T \Phi] \\ &\qquad \Phi^T A^T u = \Phi^T c \end{aligned}}$$

$$\left[\begin{aligned} &\max_u \quad u^T r \\ &\text{s.t.} \quad \Phi^T A^T u = \Phi^T c \\ &\qquad\quad u \geq 0 \end{aligned}\right]$$