solution of SDE and SOCP

goal: characterize asymptotic dynamics of SDE & cost of SOCP

solution of SDE

- consider the dynamics of SDE (X, U, P) with finite states and actions:

  $|X|, |U| < \infty, \quad x^+ \sim P(x, u)$

→ starting with an initial state distribution $p \in \Delta(X)$ and control policy $\pi : X \to \Delta(X)$, compute the next state distribution $p^+ \in \Delta(X)$

– noting that $p^+ : X \to [0,1]$ is a function, we can compute $p^+(x^+)$ for each $x^+ \in X$

$$p^+(x^+) = \sum_{x \in X} p(x) \sum_{u \in U} \pi(u|x) \cdot P(x^+ | x, u)$$

- this determines a (deterministic!) DE on the set of state distributions:

$$p^+ = F(p), \quad p \in [0,1]^X$$

→ look carefully at the definition of F:

- what kind of equation is this?
- show that $p^+ = p \cdot \Gamma$ (find Γ; what is its shape?)

- the DE is linear in p!

- $[\Gamma]_{x^+,x} = \sum_{u \in U} \pi(u|x) P(x|u,x)$,

$\Gamma \in \mathbb{R}^{N \times N}, \quad N = |X|$

* we can use linear systems theory to analyze asymptotic behavior of discrete-time linear time-invariant DE (DT-LTI) $p^+ = \Gamma p$, since $p_t = \Gamma^t p_0$!

- noting that Γ is right-stochastic, i.e.

- noting that Γ is right-stochastic, i.e. $[\Gamma]_{x_+,x} \geq 0$ and $\Gamma \cdot \underline{1} = \underline{1}$ where $\underline{1}^T = (1, \cdots, 1)$, conclude that $\forall \lambda \in \operatorname{spec} \Gamma: |\lambda| \leq 1$, i.e. the spectral radius $\rho(\Gamma) = 1$.
- if $[\Gamma]_{x_+,x} > 0$ (more generally, if Γ is irreducible & aperiodic)

  then $\bar{p} = \lim_{t\to\infty} [P^t]_{:,j}$, i.e. j-th column, is unique right-eigenvector with unity eigenvalue:

  $$\Gamma \cdot \bar{p} = \bar{p};$$

  all initial probability distributions p_0 tend to $\bar{p}$ asymptotically: $\bar{p} = \lim_{t\to\infty} P^t \cdot p_0$

$\boxed{\text{solution of SOCP}}$

- consider SOCP/MDP (X, U, P, c) with infinite-horizon exponentially-discounted cost,

  $$\min E\left[\sum_{t=0}^{\infty} \gamma^t \cdot \mathcal{J}(x_t, u_t)\right] = c(x, u)$$

$$\min \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t \cdot \mathcal{L}(x_t, u_t)\right] = c(x, u)$$

s.t. $x^+ \sim P(x, u)$

- given policy $\pi : X \to \Delta(X)$, define associated <u>value function</u> $v^\pi : X \to \mathbb{R}$

$$\forall x \in X : v^\pi(x) = \mathbb{E}[c \mid x_0 = x]$$

→ show that v^π satisfies the Bellman eq

$$v^\pi(x) = \sum_{u \in \mathcal{U}} \pi(u|x) \sum_{x^+ \in X} P(x^+|x, u) \cdot \left[\mathcal{L}(x, u) + \gamma \cdot v^\pi(x^+)\right]$$

- (this follows by pulling out first term in the sum in $\bar{c}$, re-indexing the remaining terms, and marginalizing over $u \in \mathcal{U}$ and $x^+ \in X$)

* value $v^\pi \in \mathbb{R}^X$ appears <u>linearly</u>!

→ determine L, b such that $L \cdot v^\pi = b$ (what is the shape of L, b?)

* conclude that the value of any policy can be computed by solving a linear equation?