Model-predictive control

$$\max_{\rho} V^{\rho}(x) = \mathbb{E}\left[\sum_{k=0}^{\infty} r(X_k, U_k) \,\Big|\, X_0 = x\right]$$

s.t. $\quad X_{k+1} \sim f(x_{k+1} \mid x_k, u_k)$

$\quad U_k = \rho(X_k)$

$$\downarrow$$

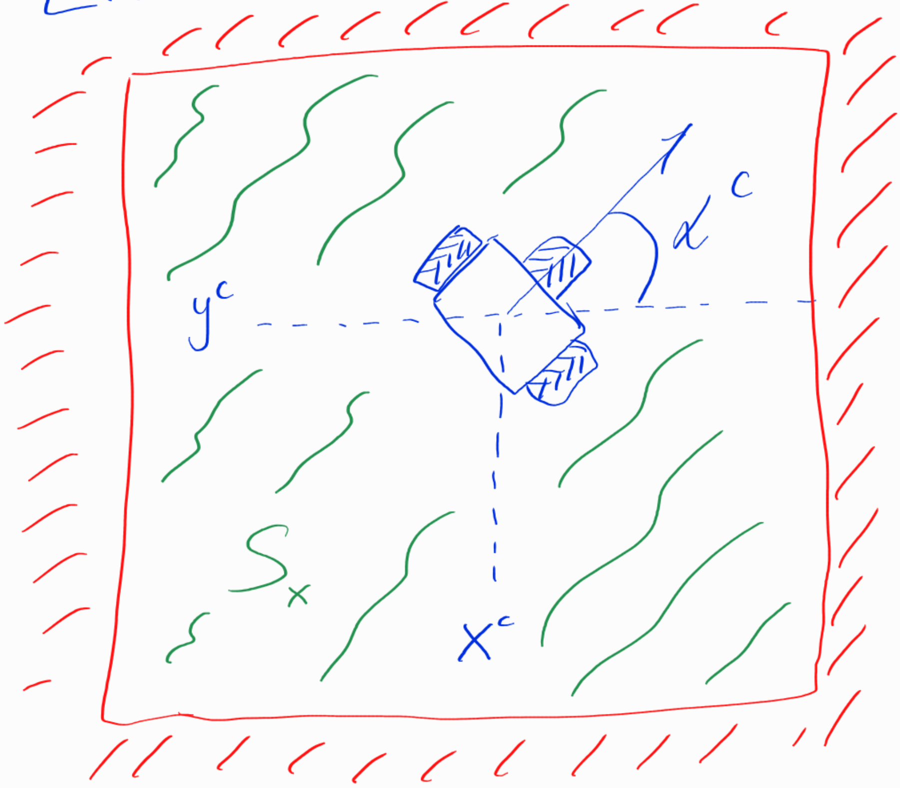
$$\max_{\{u_k\}_{k=0}^{N-1}} V_N(x) = \mathbb{E}\left[\sum_{k=0}^{N-1} r(X_k, u_k) \,\Big|\, X_0 = x\right]$$

s.t. $\quad \mathbb{P}[X_k \in S_x] \geq 1 - \varepsilon$

$\quad \mathbb{P}[U_k \in S_u] \geq 1 - \varepsilon,$

$\quad k \in \underbrace{[0 : N-1]}_{\{0, 1, \dots, N-1\}}$

Example: mobile robot



$$\mathbb{X} = \mathbb{R}^2 \times \mathbb{S}^1$$

Cart. coords $\quad$ angle

$$S_x = \{ (x^c, y^c) \in \mathbb{R}^2 : |x^c| \le \bar{x}^c, \; |y^c| \le \bar{y}^c \}$$

$$S_x \subset \mathbb{X}$$

What's about the DPP?

Recall:

$$V_N^*(x) = \max_{\{u_k\}_{k=0}^{N-1}} V_N(x)$$

$$V_N^*(x) = \max_u \{ r(x,u) + \mathbb{E}[V_{N-1}^*(X_+^u)]$$

If you take $\ell = 1$, then the identity is trivial.

Let's take $\ell = 2$:

$$V_1^*(x) = \max_{u_0} r(x, u_0)$$

$$V_2^*(x) = \max_{\{u_0, u_1\}} \left\{ r(x, u_0) + \mathbb{E}\left[r(X_+^{u_0}, u_1) \middle| X = x \right] \right.$$

Since V_1^* is optimal,

$$\max_{\{u_0, u_1\}} \left\{ r(x, u_0) + \mathbb{E}\left[r(X_+^{u_0}, u_1) \middle| X = x \right] \right. \leqslant$$

$$\max_{u_0} \left\{ r(x, u_0) + \mathbb{E}\left[V_1(X_+^{u_0}) \middle| X = x \right] \right.$$

On the other hand, by optimality of

$$u_0^*, u_1^* := \arg\max_{\{u_0, u_1\}} \left\{ r(x, u_0) + \mathbb{E}\left[r(X_+^{u_0}, u_1) \middle| X = x \right] \right.$$

no matter what tail action we take — be it the maximizer of V_1 — it must hold that

$$\max_{\{u_0,u_1\}} \left\{ r(x,u_0) + \mathbb{E}\left[r(X_+^{u_0}, u_1) \mid X = x\right] \geq \right.$$

$$\max_{u_0} \left\{ r(x,u_0) + \mathbb{E}\left[V_1(X_+^{u_0}) \mid X = x \right]\right.$$

$\Rightarrow$ identity follows

Suppose the identity holds for $l-1$, $l < N$,

$$V_{l-1}^*(x) = \max_u \left\{ r(x,u) + \mathbb{E}\left[V_{l-2}^*(X_+^u) \right]\right.$$

Unwrap the next one:

$$V_l^*(x) = \max_{\{u_k\}_{k=0}^{l-1}} \mathbb{E}\left[\sum_{k=0}^{l-1} r(X_k^{u_{k-1}}, u_k) \mid X_0 = x \right]$$

$$X_0 \xrightarrow{u_0} X_1 \xrightarrow{u_1} X_2$$

Take $k=1$, then it's $X_1^{u_0}$

$$\max_{\{u_k\}_{k=0}^{l-1}} \mathbb{E}\left[\sum_{k=0}^{l-1} r(X_k^{u_{k-1}}, u_k) \mid X_0 = x \right] =$$

$$\max_{\{u_k\}_{k=0}^{\ell-1}} \left\{ r(x, u_0) + \mathbb{E}\left[\sum_{k=1}^{\ell-1} r(X_k^{u_{k-1}}, u_k) \right] \right\}$$

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$$

tower id?

$$X_1^{u_0} \sim f(\ldots)$$
$$X_2^{u_1} \sim f(\ldots)$$

Applying the same optimality argument (as in the case of V_1^*, V_2^*), observing that

$$V_{\ell-1}^*(X_1^{u_0}) = \max_{\{u_k\}_{k=1}^{\ell-2}} \mathbb{E}\left[\sum_{k=1}^{\ell-2} r(X_k^{u_{k-1}}, u_k) \,\middle|\, X_1^{u_0} \right]$$

Observing this + optimality = proof by induction

Corollary of this is a DPP formulation with a nontrivial head:

$$V_N^*(x) = \max_{\{u_k\}_{k=0}^{N-1}} \left\{ \mathbb{E}\left[\sum_{k=0}^{N-\ell} r(X_k^{u_{k-1}}, u_k) + V_\ell^*(X_{N-\ell}) \right] \right\}$$

Original DPP (same token):

$$V^*(x) = \max_{\{u_k\}_{k=0}^{N-1}} \left\{ \mathbb{E}\left[\sum_{k=0}^{N-\ell} r(X_k^{u_{k-1}}, u_k) + V^*(X_{N-\ell}) \right] \right\}$$

$$\max_{\{u_i\}_{i=k}^{k+N-1}} V_N(x_k) = \sum_{i=k}^{k+N-1} r(x_i, u_i)$$

s.t. $x_i \in S_x, \ \forall i \in [k : k+N-1]$

$u_i \in S_u, \ \forall i \in [k : k+N-1]$

$x_{k+1} = f(x_k, u_k)$

$x_{k+N-1} = 0$ ← terminal constraint

Assumption: r is negative definite, i.e.,

$r(0,0)=0,$

$r(x,u)<0, \quad x\neq 0, u\neq 0$

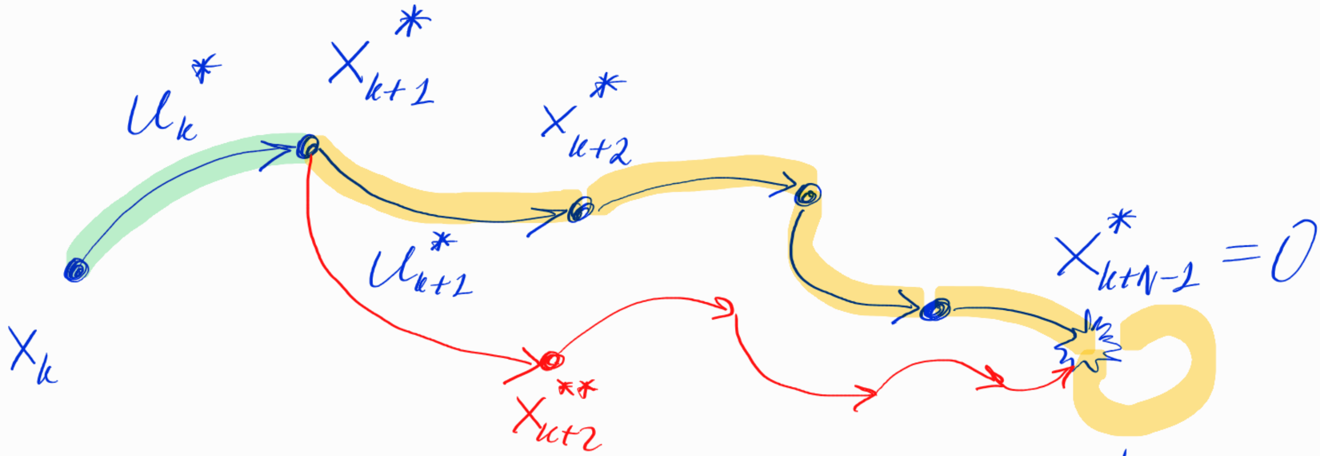
Can we guarantee that the environment achieves the target state $x=0$?

Assumption: $f(0,0)=0$

We call an action sequence $\{u_i\}_{i=k}^{k+N-1}$ feasible if it satisfies the state and actions constraints of the original problem and also the terminal constraint

Let's assume at kth step, there was a feasible solution

$\{u_i^*\}_{i=k}^{k+N-1}$



Now, we are at the step $k+1$.
Take a candidate solution

$$\{u_i\}_{i=k+1}^{k+N} = \{u_i^\#\}_{i=k+1}^{k+N} =$$

$$\{\{u_i^*\}_{i=k+1}^{k+N-1}, 0\}$$

$$\boxed{\begin{aligned} &\{u_i^*\}_{i=k}^{k+N-1} = \{u_k^*, u_{k+1}, \dots, u_{k+N-2}^*\} \\ &\{u_i^\#\}_{i=k+1}^{k+N} = \{u_{k+1}^*, u_{k+2}^*, \dots, u_{k+N-1}^*, 0\} \end{aligned}}$$

Remark: when a sequence $\{u_i^*\}_{i=k}^{k+N-1}$ was computed, the first action is actually applied to the environment. At $k+1$, the process is redone.

But, $\{u_i^{\#}\}_{i=k+1}^{k+N}$ is just a feasible candidate. In general, if we redo optimization at step $k+1$, we get some new optimal action sequence $\{u_i^{**}\}_{i=k+1}^{k+N}$.

Let's compare the optimal value from step k to step $k+1$:

$$V_N^*(x_{k+1}) - V_N^*(x_k)$$

$$V_N^*(x_k) = \sum_{i=k}^{k+N-1} r(x_i^*, u_i^*)$$

$$V_N^*(x_{k+1}) = \sum_{i=k+1}^{k+N} r(x_i^{**}, u_i^{**})$$

So,

$$\sum_{i=k+1}^{k+N} r(x_i^{**}, u_i^{**}) - \sum_{i=k}^{k+N-1} r(x_i^*, u_i^*) \geq$$

$$\sum_{i=k+1}^{k+N} r(x_i^{\#}, u_i^{\#}) - \sum_{i=k}^{k+N-1} r(x_i^*, u_i^*) =$$

$$\sum_{i=k+1}^{k+N-1} r(x_i^*, u_i^*) - \sum_{i=k}^{k+N-1} r(x_i^*, u_i^*) \geq$$

$$-r(x_k^*, u_k^*) > 0 \Leftarrow x_k^* \neq 0, \; u_k^* \neq 0$$

$\Rightarrow V_N^*$ is growing up to zero, where x gets zero

$$x = \begin{pmatrix} x^c \\ y^c \\ \alpha^c \end{pmatrix}$$

$$\dot{x} = \begin{pmatrix} v \cos\alpha^c \\ v \sin\alpha^c \\ \omega \end{pmatrix}, \quad u = \begin{pmatrix} v \\ \omega \end{pmatrix}$$