

深層強化学習

pollenJP

December 12, 2018

目次

強化学習

最適制御理論 -
Optimal Control
Theory -
動的計画法 -
Dynamic
Programming -

① 強化学習

最適制御理論 - Optimal Control Theory -
動的計画法 - Dynamic Programming -

はじめに

強化学習

最適制御理論 -
Optimal Control
Theory -
動的計画法 -
Dynamic
Programming -

今回、参考にした書籍を載せておきます。

Outline

強化学習

最適制御理論 -
Optimal Control
Theory -
動的計画法 -
Dynamic
Programming -

① 強化学習

最適制御理論 - Optimal Control Theory -
動的計画法 - Dynamic Programming -

Outline

強化学習

最適制御理論 -
Optimal Control
Theory -
動的計画法 -
Dynamic
Programming -

① 強化学習

最適制御理論 - Optimal Control Theory -
動的計画法 - Dynamic Programming -

最適制御理論 - Optimal Control Theory -

- 制御対象の状態を評価する目的関数の最大化・最小化によって制御

$$\text{目的関数 : } \int_0^T \text{cost}(x_t, u(x_t)) dt$$

$$\text{制約条件 : } x'_t = g(x_t, u(x_t))$$

T : 制御の終了時刻

x_t : 制御対象の状態変数 :

$u(x)$: 状態 x における制御命令

$\text{cost}(x)$: 状態 x において制御命令 u を実行するコスト

- この問題を解くことによってえられる $u^*(x)$ を「最適な制御則」という。 $u^*(x)$ に従う制御対象の奇跡 x_t^* を「最適経路」という。

Outline

強化学習

最適制御理論 -
Optimal Control
Theory -

動的計画法 -
Dynamic
Programming -

① 強化学習

最適制御理論 - Optimal Control Theory -

動的計画法 - Dynamic Programming -

動的計画法 - Dynamic Programming -

しかし、先の問題を解析的に解くことは不可能なことが多いため、動的計画法などを用いて数値解析的に解くことが多い。

動的計画法 最適化問題を部分問題に分割し、各部分問題の解を求めて合わせることで元の問題の解を得るというアルゴリズム。

最適性原理 Principal of Optimality.

任意の初期状態に関して決定した最適な制御則 $u^*(x)$ は、最適経路 x_t^* 上における残りの期間においても最適でなくてはならない。

Richard Bellman(1920-1984 年).

動的計画法 - Dynamic Programming -

最適制御問題を離散時間 (タイムステップ $\Delta t = 1$) で最小コスト関数 J^* 以下のように表す。

$$J_0^*(x_0) = \min \sum_{t=0}^T \text{cost}(x_t, u(x_t))$$

$J_0^*(x_0)$ 状態 x_0 に関する時刻 0 から終了時刻 T までの最小コストを出力する関数。

$J_t^*(x_t)$ 状態 x_t に関する時刻 t から終了時刻 T までの最小コストを出力する関数。

最適性原理より

$$\begin{aligned} J_0^*(x_0) &= \min \left\{ \text{cost}(x_0, u(x_0)) + \sum_{t=1}^T \text{cost}(x_t, u(x_t)) \right\} \\ &= \min \left\{ \text{cost}(x_0, u(x_0)) + \min \sum_{t=1}^T \text{cost}(x_t, u(x_t)) \right\} \\ &= \min \{ \text{cost}(x_0, u(x_0)) + J_1^*(x_1) \} \end{aligned}$$