AerE 5740: Optimal Control: Spring 2026

Iowa State University

Quiz 20: Bellman Operator in Discrete Time Infinite Horizon DP

Name:____

For each of the following multiple choice problems, ONLY ONE option is correct. Clearly mark the correct option. Calculators or other electronic devices are NOT allowed.

We mentioned in last lecture that the Bellman operator

$$\mathcal{T}V_{\infty}(\boldsymbol{x}) := \underset{\boldsymbol{u}\in\mathcal{U}}{\text{minimum}}\ \mathbb{E}\left[c(\boldsymbol{x},\boldsymbol{u},\boldsymbol{w}_k) + \beta V_{\infty}\left(\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u},\boldsymbol{w}_k)\right)\right] \quad \text{for stochastic DP,}$$

is a monotone operator in the sense

$$V_{\infty}(\boldsymbol{x}) \leq \widetilde{V}_{\infty}(\boldsymbol{x}) \quad \forall \boldsymbol{x}\in\mathcal{X} \quad \Rightarrow \quad \mathcal{T}V_{\infty}(\boldsymbol{x}) \leq \mathcal{T}\widetilde{V}_{\infty}(\boldsymbol{x}) \quad \forall \boldsymbol{x}\in\mathcal{X}.$$

All questions below are about the following three line proof for this monotonicity and its implication.

Three line proof.

$$\begin{aligned}
&V_{\infty}(\boldsymbol{x}) \leq \widetilde{V}_{\infty}(\boldsymbol{x})\\
\Rightarrow\ &\mathbb{E}\left[c(\boldsymbol{x},\boldsymbol{u},\boldsymbol{w}_k) + \beta V_{\infty}\left(\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u},\boldsymbol{w}_k)\right)\right] \leq \mathbb{E}\left[c(\boldsymbol{x},\boldsymbol{u},\boldsymbol{w}_k) + \beta \widetilde{V}_{\infty}\left(\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u},\boldsymbol{w}_k)\right)\right]\\
\Rightarrow\ &\underset{\boldsymbol{u}\in\mathcal{U}}{\text{minimum}}\ \mathbb{E}\left[c(\boldsymbol{x},\boldsymbol{u},\boldsymbol{w}_k) + \beta V_{\infty}\left(\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u},\boldsymbol{w}_k)\right)\right] \leq \underset{\boldsymbol{u}\in\mathcal{U}}{\text{minimum}}\ \mathbb{E}\left[c(\boldsymbol{x},\boldsymbol{u},\boldsymbol{w}_k) + \beta \widetilde{V}_{\infty}\left(\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u},\boldsymbol{w}_k)\right)\right].
\end{aligned}$$

Problem 1

The second line in the proof takes the expectation operator $\mathbb{E}$ to both sides of an inequality. This $\mathbb{E}$ is with respect to

(A) the noise $\boldsymbol{w}_k$ while keeping $\boldsymbol{x}, \boldsymbol{u}$ fixed. `Reason:` only the process noise $\boldsymbol{w}_k$ is random, the $\boldsymbol{x}, \boldsymbol{u}$ are symbolic parameters in DP equation.

(B) the state $\boldsymbol{x}$ while keeping $\boldsymbol{u}, \boldsymbol{w}_k$ fixed.

(C) the control input $\boldsymbol{u}$ while keeping $\boldsymbol{x}, \boldsymbol{w}_k$ fixed.

Problem 2

The reason why here $\mathbb{E}$ preserves inequality is that the operator $\mathbb{E}$ is, by definition, a weighted average with

(A) nonnegative weights. `Reason:` probability of any event is ≥ 0.

(B) strictly positive weights.

(B) strictly negative weights.

Problem 3

The last line in the proof uses that $\underset{\boldsymbol{u}\in\mathcal{U}}{\text{minimum}}$ preserves the inequality. This is because

(A) if a function is below another, then its lowest value is below the other's lowest value. `Reason:` the inequality in the second line is valid for any $\boldsymbol{x}, \boldsymbol{u}$.

(B) both functions being minimized are convex in $\boldsymbol{u}$.

(C) both functions being minimized are linear in $\boldsymbol{u}$.

Problem 4

The monotonicity proved here is used by which algorithm for discrete time infinite horizon DP?

(A) Value iteration.

(B) Policy iteration. `Reason:` Lec. 21 p. 3.

(C) Neither value nor policy iteration.

Problem 5

For discrete time infinite horizon DP with Markov decision process (MDP), the monotonicity and its proof technique

(A) generalize. `Reason:` nothing changes in the proof logic.

(B) break down.