TSRT08 Optimal Control
Homework 2 - Reinforcement Learning

Eric Moringe (erimo668), David Wiman (davwi279)

November 30, 2023

Exercises

11.9 a) Let
$$T_{Q_\mu}(Q) = f(x,u) + \gamma Q(F(x,u), \mu(F(x,u)))$$
be the Bellman Q policy operator. We want to show that if
$$Q_1(x,u) \leq Q_2(x,u) \text{ for all } (x,u),$$
then
$$T_{Q_\mu}(Q_1) - T_{Q_\mu}(Q_2) \leq 0.$$

$$\begin{aligned}
Q_1(x,u) &\leq Q_2(x,u)\\
\implies Q_1(F(x,u), \mu((F(x,u))) &\leq Q_2((F(x,u), \mu((F(x,u)))\\
\implies \gamma Q_1(F(x,u), \mu((F(x,u))) &\leq \gamma Q_2((F(x,u), \mu((F(x,u)))\\
\implies f(x,u) + \gamma Q_1(F(x,u), \mu((F(x,u))) &\leq f(x,u) + \gamma Q_2((F(x,u), \mu((F(x,u)))\\
\implies T_{Q_\mu}(Q_1) &\leq T_{Q_\mu}(Q_2)\\
\implies T_{Q_\mu}(Q_1) - T_{Q_\mu}(Q_2) &\leq 0
\end{aligned}$$

QED

We also want to show that T_{Q_μ} is a contraction. We remind ourselves of the definition. A function f is a contraction if
$$\|f(x) - f(y)\|_\infty \leq k\|x - y\|_\infty \text{ for some } 0 \leq k < 1.$$
Therefore, we want to show that
$$\|T_{Q_\mu}(Q_1) - T_{Q_\mu}(Q_2)\|_\infty \leq k\|Q_1 - Q_2\|_\infty \text{ for some } 0 \leq k < 1.$$
We rewrite the left hand side,
$$\begin{aligned}
&\|T_{Q_\mu}(Q_1) - T_{Q_\mu}(Q_2)\|_\infty\\
&= \|f(x,u) + \gamma Q_1(F(x,u), \mu(F(x,u))) - f(x,u) + \gamma Q_2(F(x,u), \mu(F(x,u)))\|_\infty\\
&= \|\gamma Q_1(F(x,u), \mu(F(x,u))) - \gamma Q_2(F(x,u), \mu(F(x,u)))\|_\infty\\
&= \gamma\|Q_1(F(x,u), \mu(F(x,u))) - Q_2(F(x,u), \mu(F(x,u)))\|_\infty
\end{aligned}$$
This gives us, by properties of the supremum norm, that
$$\gamma\|Q_1(F(x,u), \mu(F(x,u))) - Q_2(F(x,u), \mu(F(x,u)))\|_\infty \leq \gamma\|Q_1 - Q_2\|_\infty$$
For any choice of $0 \leq \gamma < 1$, it is possible to find a $0 \leq k < 1$ such that
$$\gamma\|Q_1 - Q_2\|_\infty \leq k\|Q_1 - Q_2\|_\infty$$
QED