Note

1 Some critical conditions for SSPP

A stationary policy μ is said to be proper if using that policy there is a positive probability the termination state/absorbing state will be reached.(which means eventually the termination state will be reached). In the problem done in class, if the gambler always play fair coin, the game will never end, so that is a improper policy. All other policies are proper.

1. Bellman equation is only valid for proper policies.

2. See assumption 2.1 from book: Suppose there exists at least one proper policy. Then under those proper policies you can write the bellman equation.

3. When there are no proper policy in the problem (in other words assumption 2.1 doesnot satisfy) , the optimal policy is a improper policy.

4. If you write bellman equation under a improper policy, you will see that there exists no solution of the system of equations. Which means either $J^\mu(i) = \infty$ or $J^\mu(i) = -\infty$.

5. Suppose we have a improper policy μ. If under policy μ the expected single stage cost $\sum_{j=0}^{n} p_{ij}(i)g(i, u, j) > 0$ for every non-terminal state i, then $J^\mu(i) = \infty$. If under policy μ the expected single stage cost $\sum_{j=0}^{n} p_{ij}(i)g(i, u, j) < 0$ for every non-terminal state i, then $J^\mu(i) = -\infty$.

6. See assumption 2.2 from book: For every improper policy, $J^\mu(i)$ is ∞ for some i(for a minimiztion problem as given in book). For every improper policy, $J^\mu(i)$ is $-\infty$ for some i(for a maximization problem). This condition basically means that all the improper policies doesnot have a chance of being the optimal policy (in other words the optimal policy can only be a proper policy).

7. If assumption 2.2 does not satisfy, the optimal policy is a improper policy.

8. if Assumption 2.1 or 2.2 doesnot hold, the optimal policy is a improper policy. In that case check condition 5. For a minimization problem if $J^\mu(i) = -\infty$, that is the optimal policy. For a maximization problem if $J^\mu(i) = \infty$, that is the optimal policy.

2 Bellman Operator for SSPP

Suppose J is a vector which corresponds the value function for each state, with $J(i)$ being the i-th component for state i. The dimention of the vector J is same as the size of the state space $|S| = n$. T and T^μ is the bellman operator such that if you give a vector J as input it gives:

$$TJ(i) = \min_{u \in U(i)} \sum_{j=0}^{n} p_{ij}(u)(g(i,u,j) + J(j))$$

$$T^\mu J(i) = \sum_{j=0}^{n} p_{ij}(\mu(i))(g(i,\mu(i),j) + J(j))$$

Our goal is to find J^* which satisfies the bellman equation:

$$J^* = TJ^*$$

which means we have to find a J^* which satisfies:

$$J^*(i) = \min_{u \in U(i)} \sum_{j=0}^{n} p_{ij}(u)(g(i,u,j) + J^*(j)), i = 1, \ldots, n$$

The way to solve this is to find all possible policies μ. For every policy μ solve the equation:

$$J^\mu = T^\mu J^\mu$$

which means for every policy μ find a J^μ which satisfies:

$$J^\mu(i) = \sum_{j=0}^{n} p_{ij}(a = \mu(i))(g(i, a = \mu(i), j) + J^\mu(j)), i = 1, \ldots, n$$

Then find the optimal policy which has the minimum $J^\mu(i) \forall i$