**4.**

1. The Bellman equations for the value function , :

For terminal state -

For terminal state -

1. The Bellman operators :
2. Without the assumption of all stationary policies being proper, for some improper policy we get - .

Since there is zero probability of reaching the terminal state, therefore the sum is infinite.

Also, there is no discount factor to obtain convergence of the infinite sum.

1. Proof –

Let’s consider a new SSP with same transitions and costs all equal -1, except for the terminal state, 0. Let be the optimal value from state in the new SSP and- .

Since all costs equal -1 and cost of terminal state is 0, all values from will uphold-

The optimal value also upholds the inequality above, therefore –

**a).** The Bellman equations for –

**b).** From result: , we get: . Therefore-

Returning to the original SSP problem-

Let and be two elements in , which uphold - . Let be the policy such that -

Then-

The Bellman equation for -

Therefore, for each state s –

We get-

Since , we have that is contracting with respect to the norm .

For a proper policy , we can view as the optimal Bellman operator in a new problem where the action space for each state is . For this new problem we get the desired result-

is contracting with respect to the norm .