1. $J = J^*$

lemma.

$$J(1) \le J(2) \le \cdots \le J(n)$$

Proof)

1) $J(i) \le J(i+1) \Rightarrow J(i-1) \le J(i)$

let $P_i = P_{ii}$, $K(1) = g(1) + \alpha J(1)$

there are four possible cases.

$$J(i) = \begin{cases} g(i) + \alpha P_i J(i) + \alpha(1-P_i)J(i+1) = \frac{1}{1-\alpha P_i}\left(g(i) + \alpha(1-P_i)J(i+1)\right) \cdots ① \\ g(i) + R + \alpha g(1) + \alpha^2 J(1) = g(i) + R + \alpha K(1) \cdots ② \end{cases}$$

$$J(i-1) = \begin{cases} \frac{1}{1-\alpha P_{i-1}}\left(g(i-1) + \alpha(1-P_{i-1})J(i)\right) \cdots ③ \\ g(i-1) + R + \alpha K(1) \cdots ④ \end{cases}$$

Case ①,③ (①$\le$②, ③$\le$④)

$$\left.\begin{aligned} ① &= g(i) + \alpha P_i J(i) + \alpha(1-P_i)J(i+1) \\ &\ge g(i) + \alpha J(i) \quad (\because J(i) \le J(i+1)) \\ &\ge g(i) + \alpha g(i) + \cdots = \frac{1}{1-\alpha} g(i) \Rightarrow J(i) \ge \frac{1}{1-\alpha} g(i) \end{aligned}\right\} *1$$

$$\Rightarrow J(i) \ge \frac{1}{1-\alpha} g(i-1) \quad (\because g(i) \ge g(i-1))$$

$$\Rightarrow g(i-1) \le (1-\alpha)J(i) \Rightarrow g(i-1) \le (1 - \alpha P_{i-1} - \alpha + \alpha P_{i-1})J(i)$$

$$\Rightarrow \left(g(i-1) + \alpha(1-P_{i-1})J(i)\right) \times \frac{1}{1-\alpha P_{i-1}} \le J(i)$$

$$\Rightarrow ③ \le ① \Rightarrow J(i-1) \le J(i)$$

Case ①, ④ (①$\le$②, ④$\le$③)

$$③ = \frac{1}{1-\alpha P_{i-1}}\left(g(i-1) + \alpha(1-P_{i-1})J(i)\right)$$

$$= g(i-1) + \frac{1}{1-\alpha P_{i-1}}\left(\alpha P_{i-1} g(i-1) + \alpha(1-P_{i-1})J(i)\right) \cdots *2$$

by $*1$ $\quad ② = g(i) + R + \alpha K(1) \ge ① = J(i) \ge \frac{1}{1-\alpha} g(i)$

$$\Rightarrow R + \alpha K(1) \ge \frac{\alpha}{1-\alpha} g(i)$$

$$*2 \le g(i-1) + \frac{1}{1-\alpha P_{i-1}}\left(\alpha P_{i-1} g(i) + \alpha(1-P_{i-1})(g(i) + R + \alpha K(1))\right)$$

$$= g(i-1) + \frac{1}{1-\alpha P_{i-1}}\left(\alpha g(i) + \alpha(1-P_{i-1})(R + \alpha K(1))\right)$$

$$\le g(i-1) + \frac{1}{1-\alpha P_{i-1}}\left((1-\alpha) + \alpha(1-P_{i-1})\right)(R + \alpha K(1)) \quad \left(\because R + \alpha K(1) \ge \frac{\alpha}{1-\alpha} g(i)\right)$$

$$= g(i-1) + R + \alpha K(1) = ④ \Rightarrow ③ \le ④ \quad \boxed{\text{Contradicts.}}$$

Case ②, ③

② $\geq$ ④ $\geq$ ③ $\Rightarrow J(i) \geq J(i-1)$

Case ③, ④

as $g(i) \geq g(i-1) \Rightarrow$ ③ $\geq$ ④ $\Rightarrow J(i) \geq J(i-1)$

2) $J(n) \geq J(n-1)$

$\Rightarrow$ by inserting $J(n+1) = J(n)$, (doesn't matter using virtual state $n+1$ in proof)

Prove by 1)

3) by mathematical induction, $J(i) \leq J(i-1)$ for $\forall i \in [1, n-1]$

(a) as in proof of lemma 1, case ①, ④ contradicts, optimal policy should have threshold behavior.

(b) changing $J = J_\mu$ in lemma 1, we can see that what ever the policy μ is minimization of one-step (ex - case ①, ④ should be ① $\leq$ ②, ④ $\leq$ ③ $\Rightarrow$ minimization) so that during policy iteration, threshold policy make case ①, ④ contradicts, conserved.

2.

Define new state y where $(y \in \mathbb{R}^1)$

$y_0 = c^T x_0$ and evolve with

$$\alpha_k = c^T A_k (cc^T)^{-1} c$$
$$\beta_k = c^T B_k$$
$$\gamma_k = c^T w_k$$

$$y_{k+1} = \alpha_k y_k + \beta_k u_k + \gamma_k .$$

Note that if $y_k = c^T x_k$

$$\begin{aligned} y_{k+1} &= \alpha_k y_k + \beta_k u_k + \gamma_k \\ &= c^T A_k (cc^T)^{-1} cc^T x_k + c^T B_k u_k + c^T w_k \\ &= c^T (A_k x_k + B_k u_k + w_k) = c^T x_{k+1}. \end{aligned}$$

$\Rightarrow$ By induction $y_k = c^T x_k$ for $k \in [0, N]$

with y state representation, cost of function will be

$$E_{w_k} \left[g_N(y_N) + \sum_{k=0}^{N-1} g_k(u_k) \right] \quad (u_k \text{ is action})$$

and this is one dimensional state problem