1. Give an example where Q-learning is implemented with greedy policies (i.e., \( u_t = \min_a Q_t(x_t, a) \)) and fails to converge. How can it be modified so that convergence is ensured?

2. Suppose operator \( T \) is a contraction with respect to \( \| \cdot \|_2 \). Does Gauss-Seidel value iteration converge?

3. Suppose operator \( F \) satisfies \( \|FJ - F\tilde{J}\|_2 \leq \|J - \tilde{J}\|_2 \) for all \( J, \tilde{J} \) and there is a unique \( J^* \) such that \( J^* = FJ^* \).
   
   (a) Let \( G_\gamma J = (1 - \gamma)J + \gamma FJ \). Show that there is \( \gamma \in (0,1) \) such that \( \|G_\gamma J - J^*\|_2 < \|J - J^*\|_2 \).
   
   (b) Consider \( \dot{J}_t = FJ_t - J_t \). Show that \( J_t \) converges to \( J^* \).

4. (bonus) Suppose operator \( F \) satisfies \( \|FJ - F\tilde{J}\|_\infty \leq \|J - \tilde{J}\|_\infty \) for all \( J, \tilde{J} \) and there is a unique \( J^* \) such that \( J^* = FJ^* \). Consider \( \dot{J}_t = FJ_t - J_t \). Show that \( J_t \) converges to \( J^* \).