$$\theta = A^{-1} b$$

$$A = \mathbb{E}\left[\phi_t (\phi_t - \gamma \phi_{t+1})^T\right]$$

$$b = \mathbb{E}\left[r_{t+1} \phi_t\right]$$

$$\hat{A}_t = \sum_{k=0}^{t-1} \phi_k (\phi_k - \gamma \phi_{k+1})^T + \varepsilon I$$

$$\hat{b}_t = \sum_{k=0}^{t-1} r_{t+1} \phi_t$$

$$\hat{\theta}_t = \hat{A}_t^{-1} \hat{b}_t$$

$$\hat{A}_t^{-1} = \left(\hat{A}_{t-1} + \phi_t (\phi_t - \gamma \phi_{t+1})^T\right)^{-1}$$

Sherman - Morrison formule : v_{t-1}^T

$$\hat{A}_t^{-1} = \hat{A}_{t-1}^{-1} + \frac{\hat{A}_{t-1}^{-1} \phi_t (\phi_t - \gamma \phi_{t+1})^T \hat{A}_{t-1}^{-1}}{1 + (\phi_t - \gamma \phi_{t+1}) \hat{A}_{t-1}^{-1} \phi_t}$$

$$v_{t-1} = (\hat{A}_{t-1}^{-1})^T (\phi_t - \gamma \phi_{t+1})$$

$$\hat{A}_t^{-1} = \hat{A}_{t-1}^{-1} - \frac{\hat{A}_{t-1}^{-1} \phi_t v_{t-1}^T}{1 + v_{t-1}^T \phi_t}$$

$$J(w) = \underset{\pi(w)}{V(s_0)} \qquad \nabla_w J(w) = ?$$

$$\nabla_w J_\pi(s) = \nabla_w \Big[\sum_a \pi(a,s,w)\, Q_\pi(s,a) \Big] =$$

$$= \sum_a \Big[\nabla_w \pi(a,s,w)\, Q_\pi(s,a) + \pi(s,a,w)\, \nabla_w Q_\pi(s,a) \Big] =$$

$$= \sum_a \Big[-\!\!|\!\!- \quad + \pi(a,s,w)\, \nabla_w \Big[\sum_{s'r'} p(s',r'|s,a)\,(r' + \gamma V_\pi(s')) \Big] \Big]$$

$$= \sum_a \Big[-\!\!|\!\!- + \pi(a,s,w) \sum_{s',r'} p(s',r'|s,a)\, \gamma\, \underline{\nabla_w V_\pi(s')} \Big]$$

$$= \sum_a \Big[\nabla_w \pi(a,s,w)\, Q_\pi(s,a) + \pi(a,s,w) \sum_{s',r'} p(s',r'|s,a)\, \gamma \cdot$$

$$\Big(\sum_{a'} \nabla_w \pi(a',s',w)\, Q_\pi(s',a') + \pi(a',s',w) \sum_{s'',r''} p(s'',r''|s',a')\, \gamma\, \nabla_w V_\pi(s'') \Big)$$

$$= \ldots = \sum_{x \in S} \sum_{k=0}^{\infty} \gamma^k\, p(s \to x, k, \pi) \sum_a \nabla_w \pi(a,x,w)\, Q_\pi(x,a)$$

$$\nabla_w J(w) = \nabla_w V_\pi(s_0)$$

$$= \sum_s d_\pi(s) \sum_a \nabla_w \pi(a,s,w)\, Q_\pi(s,a)$$

$$\nabla_w J(w) = E_\pi \Big[\gamma^t \sum_a \nabla_w \pi(a,s,w)\, Q_\pi(s,a) \Big]$$

$$= E_\pi \Big[\gamma^t \sum_a \pi(a,s,w) \frac{\nabla_w \pi(a,s,w)}{\pi(a,s,w)} Q_\pi(s,a) \Big]$$

$$= E_\pi \Big[\gamma^t \underline{\sum_a \pi(a,s,w)}\, \nabla_w \log \pi(a,s,w)\, Q_\pi(s_t,a) \Big)$$

$$= E_\pi \Big[\gamma^t\, \nabla_w \log \pi(a,s,w)\, Q_\pi(s,a) \Big]$$