Exercise 2: Principle Component Analysis

1. $f(GQ) = tr(Q^TG^TSGQ) = tr(Q^TQ \cdot G^TSG) = tr(G^TSG)$. $\forall$ orthogonal matrix $Q \in \mathbb{R}^{k \times k}$

2. Denote the SVD of $\tilde{X}$ by $\tilde{X} = U\Sigma V^T$. $\Sigma = diag(\lambda_1, \dots, \lambda_d)$. $\lambda_1 \geq \dots \geq \lambda_d \geq 0$.

   thus, $S = \frac{1}{n-1} U\Sigma^2U^T$. $y = U^Tg$. so, $\|y\| = 1$.

   therefore, $f(g) = \frac{1}{n-1} g^TU\Sigma^2U^Tg = \frac{1}{n-1}\sum_{i=1}^{d} y_i^2\lambda_i^2 \leq \frac{1}{n-1}\lambda_1^2$.

   note that, the equivalence condition is $U^Tg = e_1$

   i.e $g = Ue_1$. is the first principal component vector of the data.

3. Note that. the columns of U, $\{u_i = Ue_i\}_{i=1}^{d}$, are the orthonormal basis of $\mathbb{R}^d$.

   then, $\forall g \in \mathbb{R}^d$. with $\|g\| = 1$. $g = \sum_{i=1}^{d} a_iu_i$. and $\sum_{i=1}^{d} a_i^2 = 1$.

   since $\langle g_1, g \rangle = 0$, we have $\langle u_1, g \rangle = a_1 = 0$.

   so, $f(g) = \frac{1}{n-1}\sum_{i=2}^{d} a_i^2\lambda_i^2 \leq \frac{1}{n-1}\lambda_2^2$.

   the equivalence condition is $a_2 = 1$.

   i.e $g_2 = Ue_2$. the second principal component vector of the data.

4. the first k principal component vector is Ue_k.

5. $f(g_k) = \frac{1}{n-1}\lambda_k^2$

6. $\lambda_1 > \lambda_2 > \dots > \lambda_k$.

Exercise 5: Grid World with a Given Policy

1.

0.25	0.05									
0.7	0.8	0.05								
	0.15	0.25								
0.05			0.3		0.15					
		0.7		0.3						
			0.7		0.1			0.7		
					0.7	0.25	0.05			
				0.7		0.7	0.25			
					0.05			0.25		
						0.05		0.05	1	
							0.7			1

1a). $P(S_1 = s_i) = \frac{1}{11}$.

0.3
1.55
0.4
0.5
1
1.5
1
1.65
0.3
1.1
1.7

$P(S_2 = s_i) = \frac{1}{1}$

0.1525
1.47
0.3325
0.39
0.58
0.71
1.3825
1.8125
0.15
1.165
2.855

2. 1a). $P(S_1 = s_i) = \sum_{j=1}^{11} m_{ij}$. $\forall i = 1, 2, \dots, 11$.

let $u=(1,\dots,1)^T \in \mathbb{R}^{11}$. $P_i = (P(S_i=s_1)\cdots P(S_i=s_{11}))^T$. $i=1,2$.

then $P_1 = \frac{1}{11} M u$. $P_2 = M P_1 = \frac{1}{11} M^2 u$.

(1b). $P(S_t = s_i) = [\frac{1}{11} M^t u]_i = \frac{1}{11} e_i^T M^t u$. $i=1,2,\dots,9$.

(1c).

3. $V^\pi(s_1) = 0.9^4 \times 100 = 65.61$

$V^\pi(s_2) = 0.9^3 \times 100 = 72.9$

$V^\pi(s_3) = 0.9^2 \times 100 = 81$

$V^\pi(s_4) = 0.9^3 \times 100 = 72.9$

$V^\pi(s_5) = 0.9 \times 100 = 90$.

$V^\pi(s_6) = 0.9^2 \times 100 = 81$

$V^\pi(s_7) = 0.9 \times 100 = 90$.

$V^\pi(s_8) = 100$.

$V^\pi(s_9) = 0.9^3 \times 100 = 72.9$.

4. $P(S_t = s_i) = [M^t p_0]_i = e_i^T M^t p_0$. $i=1,2,\dots,9$.

Exercise 7: Value Iteration and Policy Iteration

1. $V^*(s) = \max_a Q(s,a) = \max_a r(s,a) + \gamma V^*(\delta(s,a))$.

$$Q(s,a) = r(s,a) + \gamma \max_{a'} Q(\delta(s,a), a').$$

$$V^{k+1}(s) = \max_a Q(s,a) = \max_a r(s,a) + \gamma V^k(\delta(s,a)).$$

$$|V^*(s) - V^{k+1}(s)| = \left|\max_a [r(s,a) + \gamma V^*(\delta(s,a))] - \max_a [r(s,a) + \gamma V^k(\delta(s,a))]\right|$$

$$\le \max_a \left|[r(s,a) + \gamma V^*(\delta(s,a))] - [r(s,a) + \gamma V^k(\delta(s,a))]\right|$$

$$= \gamma \max_a |V^*(\delta(s,a)) - V^k(\delta(s,a))| \qquad \forall s.$$

$$\Rightarrow \|V^* - V^{k+1}\|_\infty \le \gamma \|V^* - V^k\|_\infty$$

2.(a) $Q^\pi(s,a) = r(s_0,a_0) + \gamma r(s_1,a_1) + \gamma^2 r(s_2,a_2) + \cdots$ $\quad s_0 = s,\ a_0 = a.$ $\begin{cases} s_t = \delta(s_{t-1}, a_{t-1}) \\ a_{t+1} = \pi(s_{t+1}) \end{cases}$ $t=1,2,\dots$

$$= r(s_0,a_0) + \gamma Q^\pi(s_1,a_1)$$

$$= r(s,a) + \gamma\, Q^{\pi}(\delta(s,a), \pi(s)).$$

$$V^{\pi}(s) = R(s) + \gamma \sum_{s' \in S} P(s' \mid s, \pi(s))\, V^{\pi}(s'). \qquad V^{\pi} = R + \gamma P^{\pi} V^{\pi}.$$

1b). $$V^{\pi'}(s) = R(s) + \gamma \sum_{s' \in S} P(s' \mid s, \pi'(s))\, V^{\pi'}(s').$$

$$\pi'(s) = \operatorname*{argmax}_{a \in A} Q^{\pi}(s,a)$$

$$\begin{aligned} V^{\pi'}(s) &= R(s) + \gamma \max_{a \in A} \sum_{s' \in S} P(s' \mid s, \pi'(s))\, V^{\pi}(s') \\ &\geq R(s) + \gamma \sum_{s' \in S} P(s' \mid s, \pi'(s))\, V^{\pi}(s') \\ &= V^{\pi}(s). \end{aligned}$$