Exercise Sheet 4

1 Sparse Coding

a) J is convex $\rightarrow$ setting gradient to 0 will yield minimum

$$J = \frac{1}{N}\sum_{i=1}^{N} \|x_i - W s_i\|^2 + \lambda \cdot \frac{1}{N}\sum_{i=1}^{N} \|s_i\|_1 + \eta \|W\|_F^2$$

$$\frac{\partial J}{\partial W} = \frac{\partial}{\partial W}\frac{1}{N}\sum_{i=1}^{N} \|x_i - W s_i\|^2 + \eta \|W\|_F^2 \overset{!}{=} 0$$

$$\Leftrightarrow \frac{\partial}{\partial W}\frac{1}{N}\sum_{i=1}^{N} (x_i - W s_i)^T (x_i - W s_i) + \eta \|W\|_F^2 = 0$$

$$\Leftrightarrow \frac{1}{N}\sum_{i=1}^{N} -2(x_i - W s_i) s_i^T + 2\eta W = 0$$

$$\Leftrightarrow \frac{1}{N}\sum_{i=1}^{N} -(x_i - W s_i) s_i^T + \eta W = 0$$

$$\Leftrightarrow \frac{1}{N}\sum_{i=1}^{N} W(\eta \mathbb{1} + s_i s_i^T) = \frac{1}{N}\sum_{i=1}^{N} x_i s_i^T$$

$$\Leftrightarrow W(\Sigma_{ss} + \eta \mathbb{1}) = \Sigma_{xs} \quad \Leftrightarrow W = \Sigma_{xs}(\Sigma_{ss} + \eta \mathbb{1})^{-1}$$

b) $$\frac{\partial}{\partial s_i} J \Rightarrow \frac{\partial}{\partial s_i}\frac{1}{N}\sum_{i=1}^{N} \|x_i - W s_i\|^2 + \lambda \frac{1}{N}\sum_{i=1}^{N} q_i^T s_i \overset{!}{=} 0$$

$$\Leftrightarrow \frac{1}{N}\sum_{i=1}^{N} \frac{\partial}{\partial s_i}(x_i - W s_i)^T (x_i - W s_i) + \lambda \frac{1}{N}\sum_{i=1}^{N} \frac{\partial}{\partial s_i} q_i^T s_i = 0$$

$$\Leftrightarrow \frac{2}{N} W^T (x_i - W s_i) = \frac{\lambda}{N} q_i^T$$

$$\Leftrightarrow W^T x_i - W^T W s_i = \frac{\lambda}{2} q_i^T$$

$$\Leftrightarrow s_i = (W^T W)^{-1}(W^T x_i - \frac{\lambda}{2} q_i^T)$$

2 Auto-Encoders

a) $\underset{\alpha,u}{\text{argmin}} \frac{1}{N}\sum_{i=1}^{N} \|x_i - \hat{x}_i\|^2 = \underset{\alpha,u}{\text{argmin}} \frac{1}{N}\sum_{i=1}^{N} \|x_i - w(w^T x_i)\|^2$

$= \underset{\alpha,u}{\text{argmin}} \frac{1}{N}\sum_{i=1}^{N} \|x_i - \alpha u(\alpha u^T x_i)\|^2$

$= \underset{\alpha,u}{\text{argmin}} \frac{1}{N}\sum_{i=1}^{N} \|x_i - \alpha^2 u u^T x_i\|^2$

$= \underset{\alpha,u}{\text{argmin}} \frac{1}{N}\sum_{i=1}^{N} (x_i^T - \alpha^2 u u^T x_i)^T (x_i - \alpha^2 u u^T x_i)$

$= \underset{\alpha,u}{\text{argmin}} \frac{1}{N}\sum_{i=1}^{N} (\alpha^4 - 2\alpha^2)(u^T x_i x_i^T u)$

$= \underset{\alpha,u}{\text{argmin}} (\alpha^4 - 2\alpha^2)(u^T (\frac{1}{N}\sum_{i=1}^{N} x_i x_i^T) u)$

$= \underset{\alpha,u}{\text{argmin}} (\alpha^4 - 2\alpha^2)(u^T S u) = \underset{\alpha,u}{\text{argmax}}\ u^T S u$