Excercise - 5

ex - 1

(a) $$\underset{W, s_1 \ldots s_N}{E} = \eta \|W^2\|_F + \sum_{j=1}^{N} \|x_j - W s_j\|^2 + \lambda \|s_j\|_1$$

$$\frac{\partial E}{\partial W} = 2\eta \|W\|_F + \sum_{j=1}^{N} (-2 s_j)^{T} \|x_j - W s_j\|$$

$$= 2\eta W - 2 \sum_{j=1}^{N} (x_j - W s_j)\, s_j^{T}$$

(b) $$\frac{\partial E}{\partial s_j} = \sum_{j=1}^{N} (-2W)^{T} (x_j - W s_j) + \lambda$$

$$= \lambda - 2 \sum_{j=1}^{N} W^{T} (x_j - W s_j)$$

ex - 2

(a) $s_j = g(r_j)$ → Given — (1)

comparing the regularized terms from both the equations.

$$\forall_{j=1}^{N} : \|s_j\|_1 = \|g(r_j)\|_1 = \|r_j\|^2 \quad \text{— (2)}$$

$\vec{s}_j$ is always positive, as per constraint. $g(r_j)$ can be written

$$\Rightarrow g_j(r_j) = r_{jj}^{2}$$

$$j = [1 \ldots h]$$

2. (b) formulation of optimization objective can endow us with an initial guess which can save a lot of iterations and ~~anbemore~~ efficient. ~~to~~

However, this formulation can be dependent upon r_j, which also need to be learnt in order to find source s_j and can be proven time taking.

3.

(a) Given $r_1 = v^T x_j$

$$E = \eta \|\omega\|_F^2 + \sum_{j=1}^{N} \|x_j - W g(v^T x_j)\|^2 + \lambda \| v^T x_j \|^2$$

Now $\frac{\partial E}{\partial V}$

$$= \sum_{j=1}^{N} \left(-2 W g(v^T x_j) \right)^T (x_j - W g(v^T x_j)) + 2\lambda (v^T x_j) \|x_j\|^2$$

$$\frac{\partial E}{\partial V} = 2 \left(\lambda v^T \|x_j\|^2 - (W g(v^T x_j))^T (x_j - W g(v^T x_j)) \right)$$

3-(b) formulation of sparse coding enables to use back propagation to ~~train~~ ~~the~~ estimate the source which in normal sparse coding has to be found ~~out~~ by ~~us~~ taking the gradient.

Training the system ~~can~~ by backpropagation can be optional and save iterations. But also posit the risk of using wrong model for training which would be very in-efficient.