Machine learning Homework- Dimensionality Reduction

Abinav Ravi Venkatakrishnan - 03694216 and Abhijeet Parida - 03679676

January 20, 2019

1 PCA and SVD

Problem 1:

The model $y = Ax$ is a noiseless transformation since we haven't included the bias.Given that
$p(z) = \mathcal{N}(z|0, I)$
$p(x|z) = \mathcal{N}(x|Wz + \mu, \phi)$
we come to a conclusion that x is a gaussian distribution which gives us results that mean of the distribution is μ and the variance is $WW^T\phi$.
from this we can get the distribution of y with mean and variances $A\mu$ and $AWW^TA^T + A\phi A^T$. From pattern matching we can say that the maximum likelihood can be written as $A\mu_x$ and AW_x and $A\phi_x A^T$.

Problem 2:

The representation of the new vector in the space is given by [0,3,0,0,4].So the new Matrix M looks like $\begin{Bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 3 & 0 & 0 & 4 \end{Bmatrix}$ The SVD of the matrix gives U,V,and S which gives the singular values and vector. Now to do dimensionality reduction we take only top K singular values of the output and we get a P = M*V. as P = [0,3,0,0,4]*V = [1.74,2.84] which mostly corresponds to the classic movie genre. A good recommendation of the genre would consists of Titanic and Casablanca. Since the previous data consists of Titanic we can recommend Casablanca.

Problem 5:

(a)The Autoencoder cannot have reconstruction error to be zero since it would correspond to PCA and hence has a linear mapping between latent space and the decoder output. This is not a desirable property.
(b) This is possible only in the case of a linear mapping conditions where the entire input is learnt as such.