Exercise

12

TUM Department of Informatics

Supervised by Prof. Dr. Stephan Günnemann

Informatics 3 - Professorship of Data Mining and Analytics

Submitted by Marcel Bruckner (03674122)

Julian Hohenadel (03673879)

Kevin Bein (03707775)

Submission date Munich, January 23, 2020

Dimensionality Reduction & Clustering

Problem 1:

Problem 2:

(Linear) Autoencoder:

Input data X: D-dimensional

Hidden layer: K-dimensional

No biases, activations = identity.

This results in a linear transformation: $f(x) = f_{dec}(f_{enc}(x)) = XW_1W_2$

With dimensions: $X : N \times D, W_1 : D \times K, W_2 : K \times D$

With $K < D, XW_1$ forces X into a K dimensional5 subspace.

Since this transformation is not the identity ($K < D$) perfect reconstruction is not achievable unless the input data X is already in a K-dimensional subspace despite being D-dimensional data.

Problem 3:

K Gaussians:

Intuition: Expected value of Gaussian is the mean.

$\implies$ Expected value of K Gaussians should be the K means added up (each cluster z).

$$p(x) = \sum_k \pi_k \mathcal{N}(x|\mu_k, \Sigma_k)$$

$$\mathbb{E}[x] = \mathbb{E}_{p(z)}[\mathbb{E}_{p(x|z)}[x|z]]$$

$\mathbb{E}_{p(x|z)}[x|z]$ is the expected value of x in cluster z which is the mean of cluster z.

This implies $\mathbb{E}_{p(x|z)}[x|z] = \mu_k$.

$\mathbb{E}_{p(z)}$ is the prior probability π_z of z, also have to consider all clusters.

This implies $\mathbb{E}_{p(z)} = \sum_{k=1}^{K} \pi_k$.

Simply filling into the equation yields: $\mathbb{E}[x] = \sum_{k=1}^{K} \pi_k \mu_k$

Now $Cov[x]$:

$\mathbb{E}[x]$ and $\mathbb{E}[x]^T$ respectivly are already known.

$$Cov[x] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T$$

So only $\mathbb{E}[xx^T]$ is still missing:

Filling into $\mathbb{E}[x] = \mathbb{E}_{p(z)}[\mathbb{E}_{p(x|z)}[x|z]]$:

$$\mathbb{E}[xx^T] = \mathbb{E}_{p(z)}[\mathbb{E}_{p(x|z)}[xx^T|z]]$$
$$\mathbb{E}[xx^T] = \sum_{k=1}^{K} \pi_k \mathbb{E}_{p(x|z)}[xx^T|z]$$

With $\Sigma = \mathbb{E}[(X-\mu)(X-\mu)^T] = \mathbb{E}[XX^T] - \mu\mu^T$
$\implies \mathbb{E}[XX^T] = \Sigma + \mu\mu^T$

$$\mathbb{E}[xx^T] = \sum_{k=1}^{K} \pi_k(\Sigma_k + \mu_k\mu_k^T)$$
$$\implies Cov[x] = \sum_{k=1}^{K} \pi_k(\Sigma_k + \mu_k\mu_k^T) - \sum_{k=1}^{K}\sum_{l=1}^{K} \pi_k\pi_l\mu_k\mu_l^T$$

Problem 4:

a) According to slide 12-19, we first need to draw the cluster indicators k, l from categorical distribution given the probabilities π_k^x, π_l^y

$$k \sim Cat(\pi_k^x)$$
$$l \sim Cat(\pi_l^y)$$

and then sample from their respective normal distributions:

$$x \sim \mathcal{N}(\mu_k^x, \Sigma_k^x)$$
$$y \sim \mathcal{N}(\mu_k^x, \Sigma_k^x)$$

z is then given by $z := x + y$

b) Since x and y are i.i.d and $x \sim \mathcal{N}(\mu_k^x, \Sigma_k^x)$ and $y \sim \mathcal{N}(\mu_k^y, \Sigma_k^y)$, the sum of two Gaussian distributions is again Gaussian. Therefore $p(z \,|\, \theta^x, \theta^y)$ is a mixture of Gaussians.

c)

$$p(z \,|\, \theta^x, \theta^y) = \sum_{k=1}^{K_x}\sum_{l=1}^{K_y} \pi_k^x \pi_l^y \, \mathcal{N}(z \,|\, \mu_k^x + \mu_l^y, \Sigma_k^x + \Sigma_l^y)$$

Problem 5:

Appendix

We confirm that the submitted solution is original work and was written by us without further assistance. Appropriate credit has been given where reference has been made to the work of others.

Munich, January 23, 2020, Signature Marcel Bruckner (03674122)

Munich, January 23, 2020, Signature Julian Hohenadel (03673879)

Munich, January 23, 2020, Signature Kevin Bein (03707775)