Homework 2: Latent Variable Models

Deliverable: This PDF write-up by **Wei-Jer Chang Wed February 21st, 23:59pm**. Your PDF should be generated by simply replacing the placeholder images of this LaTeX document with the appropriate solution images that will be generated automatically when solving each question. The solution images are automatically generated and saved using the accompanying IPython notebook. Your PDF is to be submitted into Gradescope. This PDF already contains a few solution images. These images will allow you to check your own solution to ensure correctness.

Question 1: VAEs on 2D Data [20pt]

(a) [10pt] Data from a Full Covariance Gaussian

Final -ELBO: 5.4568, Recon Loss: 3.8836, KL Loss: 1.5732 (Dataset 1)

(a) Training curve (b) Samples with Noise (c) Samples without Noise

Figure 1: Results for Dataset 1

Final Full -ELBO: Final -ELBO: 5.4453, Recon Loss: 3.8504, KL Loss: 1.5949(Dataset 2)

(a) Training curve (b) Samples with Noise (c) Samples without Noise

Figure 2: Results for Dataset 2

(b) **[10pt] Data from a Diagonal Gaussian**

Final Final -ELBO: 5.4358, Recon Loss: 3.7938, KL Loss: 1.6420 (Dataset 1)

(a) Training curve (b) Samples with Noise (c) Samples without Noise

Figure 3: Results for Dataset 1

Final Final -ELBO: 5.4351, Recon Loss: 3.8261, KL Loss: 1.6091 (Dataset 2)

(a) Training curve (b) Samples with Noise (c) Samples without Noise

Figure 4: Results for Dataset 2

Answer: For a the latents are being used, (we learn a good mean of the distribution). On the other hand for (b) is ignored. This is because when the decoder is too powerful, we may ignore the latents.

Question 2: VAEs on Images [40pt]

(a) **[20pt] VAE**

Final -ELBO: 117.6243, Recon Loss: 93.3558, KL Loss: 24.2686 (Dataset 1)

(a) Training Curve

(b) Samples

(c) Reconstructions

(d) Interpolations

Figure 5: Results for Dataset 1

Final Final -ELBO: 204.2054, Recon Loss: 147.0967, KL Loss: 57.1087 (Dataset 2)

(a) Training Curve

(b) Samples

(c) Reconstructions

(d) Interpolations

Figure 6: Results for Dataset 2

(b) **[20pt] Hierarchical VAE with Learned Prior**

Final -ELBO: 120.5412, Recon Loss: 93.8273, KL Loss: 26.7139 (Dataset 1)

(a) Training Curve

(b) Samples

(c) Reconstructions

(d) Interpolations

Figure 7: Results for Dataset 1

Final -ELBO: 203.4307, Recon Loss: 153.1637, KL Loss: 50.2670 (Dataset 2)

(a) Training Curve

(b) Samples

(c) Reconstructions

(d) Interpolations

Figure 8: Results for Dataset 2

Question 3: VQ-VAE [40pt]

VQ-VAE Final Test Loss: 2160.6643(summed up latent dimension, mean over batch) Transformer Prior Final Test Loss: 1.3106 (Dataset 1)

(a) VQ-VAE Training Curve

(b) Transformer Prior Training Curve

(c) Samples

(d) Reconstructions

Figure 9: Results for Dataset 1

Final VQ-VAE Test Loss: 549.5761, PixelCNN Prior Test Los: 2.2399 (Dataset 2)

(a) VQ-VAE Training Curve

(b) PixelCNN Prior Training Curve

(c) Samples

(d) Reconstructions

Figure 10: Results for Dataset 1