

Informe Práctica:

Auto-encoding Variational Bayes

Marcos Martínez Jiménez

Informe de las pruebas realizadas para resolver la práctica y los resultados obtenidos.

1. Implementación

- Estimación del ELBO: nada que comentar, de acuerdo a las indicaciones provistas en la práctica.
- Hiperparámetros de aprendizaje en ADAM: se utilizaron como valores de los hiperparámetros α , β_1 , β_2 y ϵ aquellos indicados como buenos valores por defecto en la Figura 5 del enunciado.

2. Modelo Generativo

La Figura 1 muestra las 25 imágenes generadas mediante el modelo generativo al muestrear variables latentes de la distribución a priori y usarlas para obtener probabilidades de cada píxel.

Se puede comprobar que, como cabría esperar, las imágenes generadas parecen corresponder a dígitos concretos (Tabla 1) y son relativamente compatibles con dibujos que pudiera hacer un ser humano.

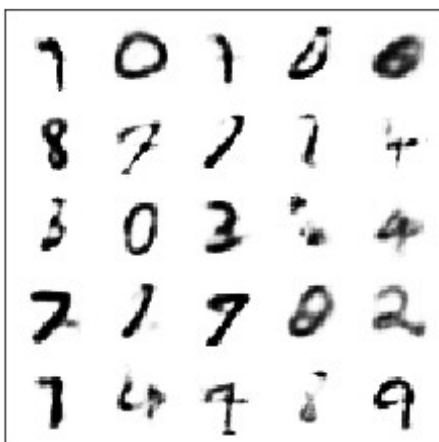


Figura 1: Imágenes generadas con el modelo generativo.

1/7	0	7	0	6
8	7/9	1/9	2/7	4
3	0	3	5	4
7	1	9	8	2
1/7	4	4/9	8	9

Tabla 1: Correspondencia entre imágenes y dígitos.

3. Reconstrucción de imágenes

La Figura 2 muestra la reconstrucción mediante el VAE de las 10 primeras imágenes de test. Para producirlas se obtuvo una muestra de la distribución posterior de variables latentes que genera el encoder, y se utilizó dicha muestra para obtener las probabilidades de cada píxel con el decoder.

Se puede observar que la reconstrucción de las imágenes es casi perfecta en la mayoría de los dígitos (la del 5 es ligeramente peor), aunque se gana algo de ruido.

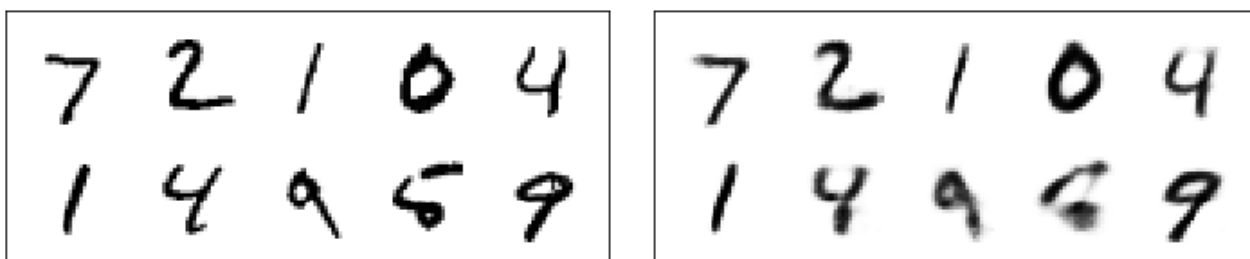


Figura 2: Reconstrucción de imágenes. Se muestran las primeras 10 imágenes del conjunto de test (*izquierda*) y sus reconstrucciones utilizando el VAE (*derecha*).

4. Interpolación en el espacio latente

La Figura 3 muestra las interpolaciones entre los 5 primeros pares de imágenes de test, que se obtuvieron de la forma indicada en el enunciado de la práctica.

En todos los casos se observan claramente los dígitos correspondientes a las imágenes reales (las de los extremos), y se puede ver como las imágenes intermedias corresponden a mezclas entre ambos (representando en muchos casos la forma de dibujar el dígito más ambigua entre los dígitos reales).

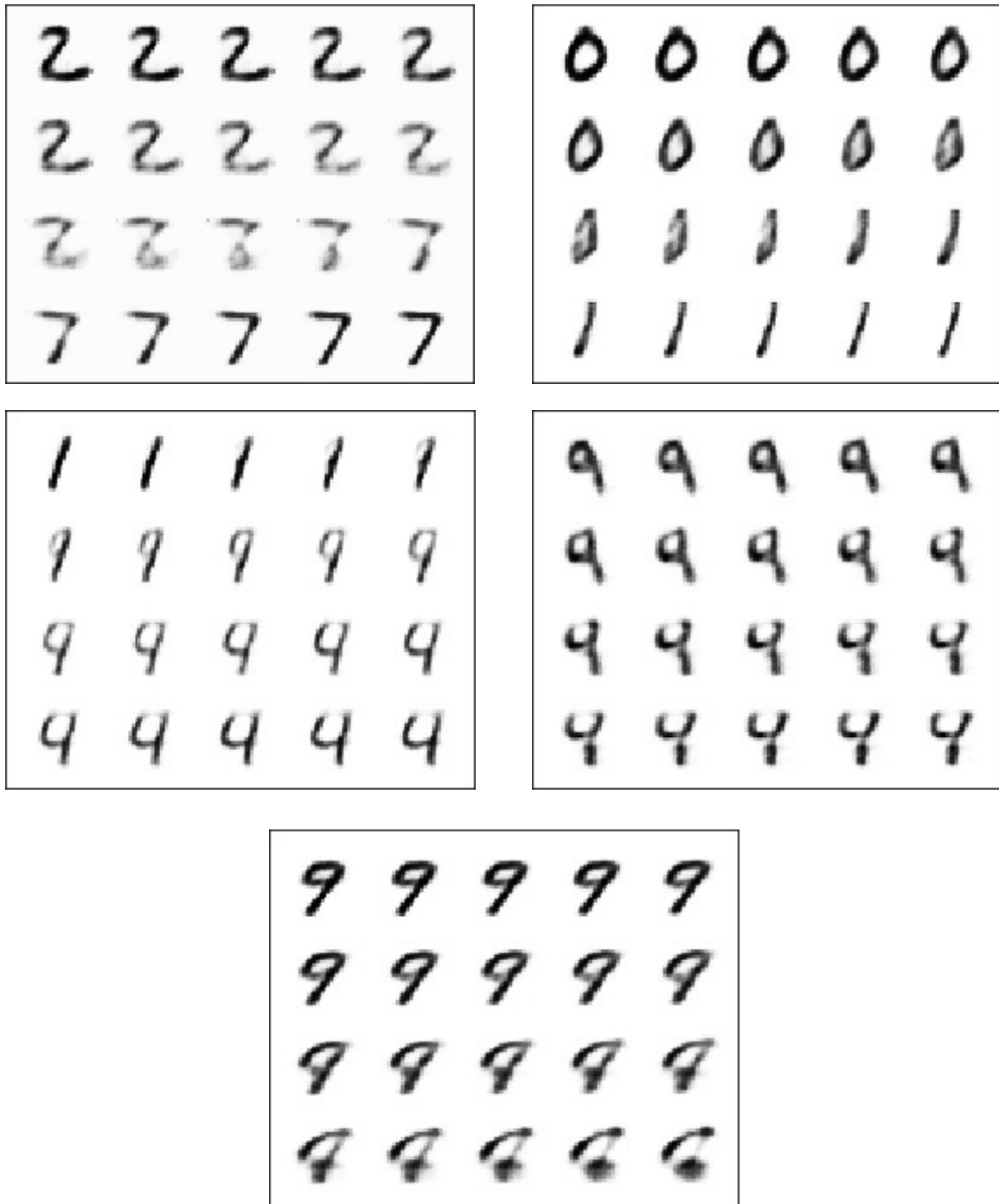


Figura 3: Interpolación en el espacio latente. Se muestran las imágenes generadas a partir de la interpolación en el espacio latente de los 5 primeros pares de imágenes de test.