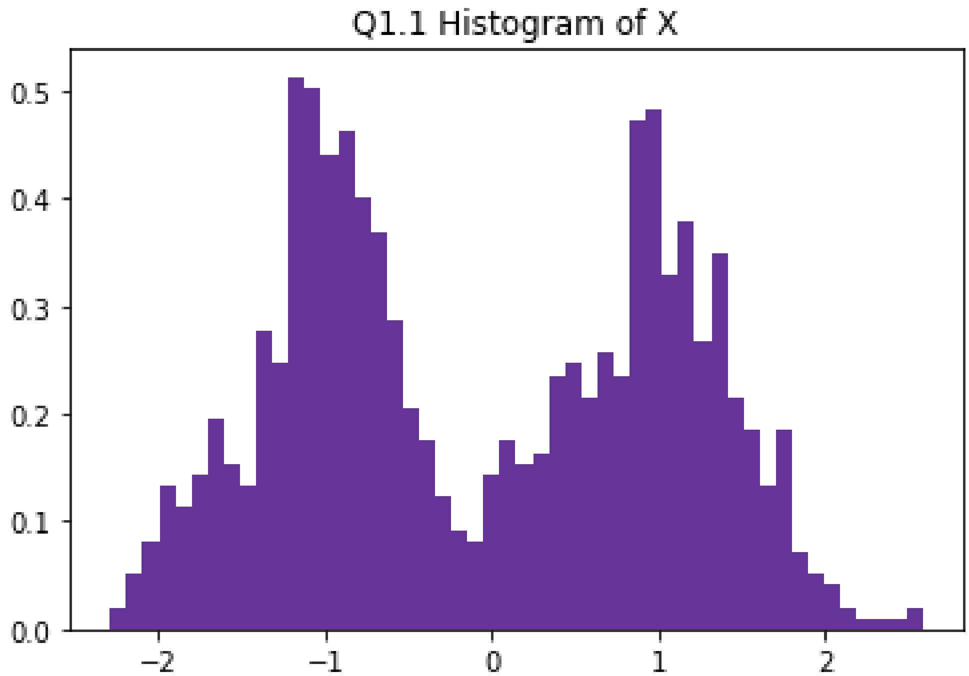
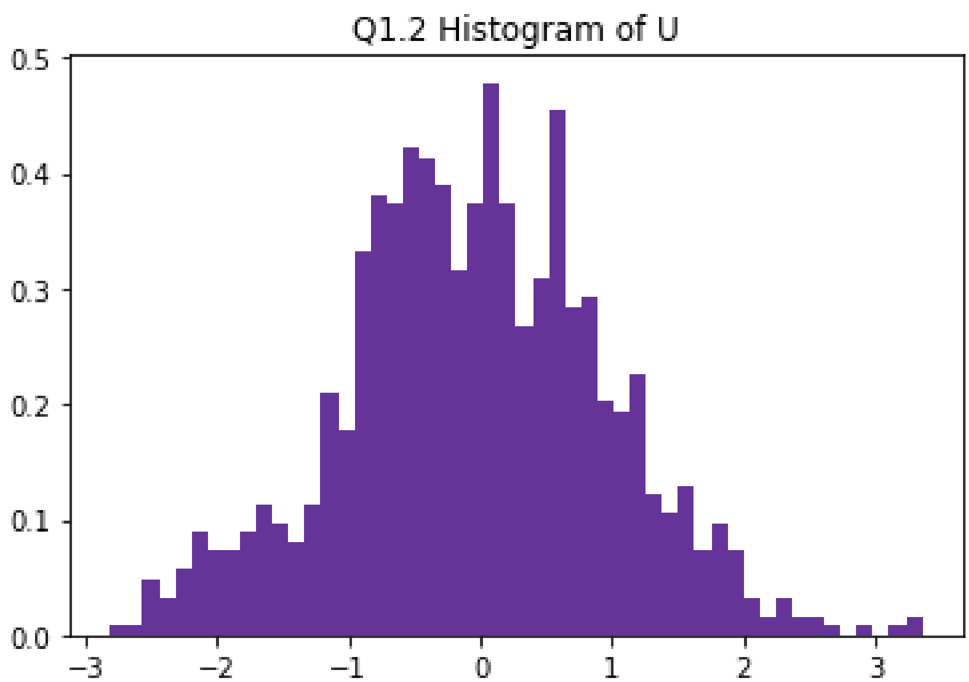
CS680 – Assignment 4

Ali ElSaid (20745892)

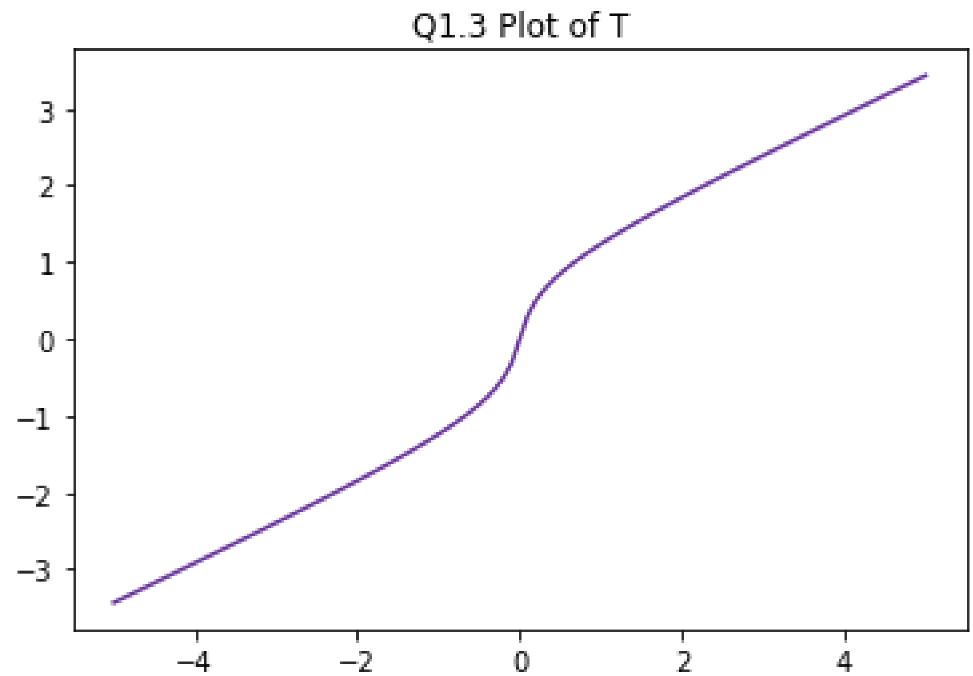
Q1.1:



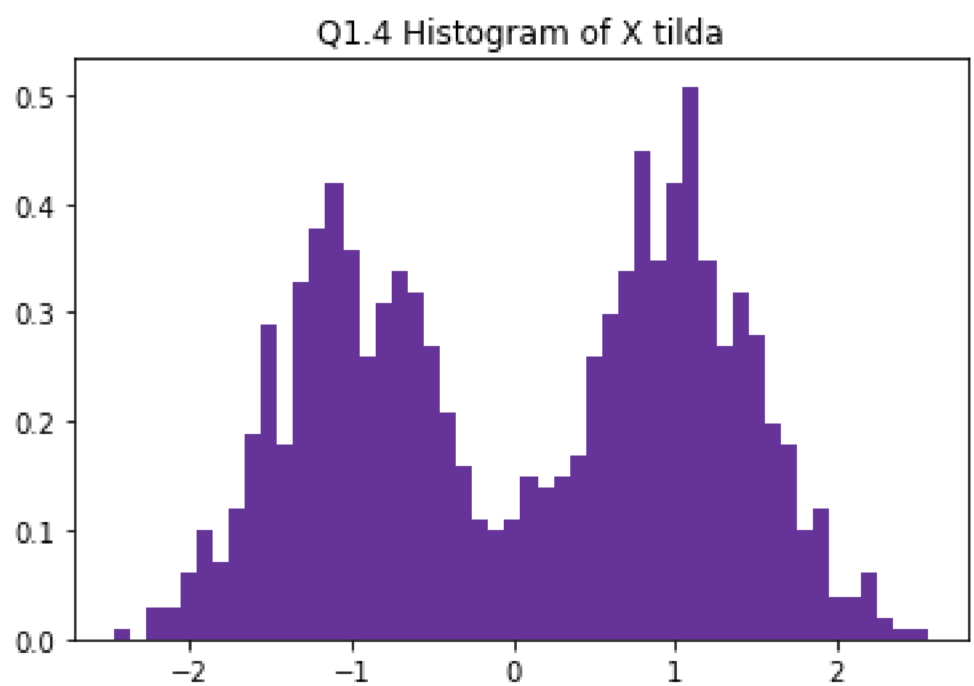
Q1.2: Ui should follow a standard Gaussian distribution approximately.



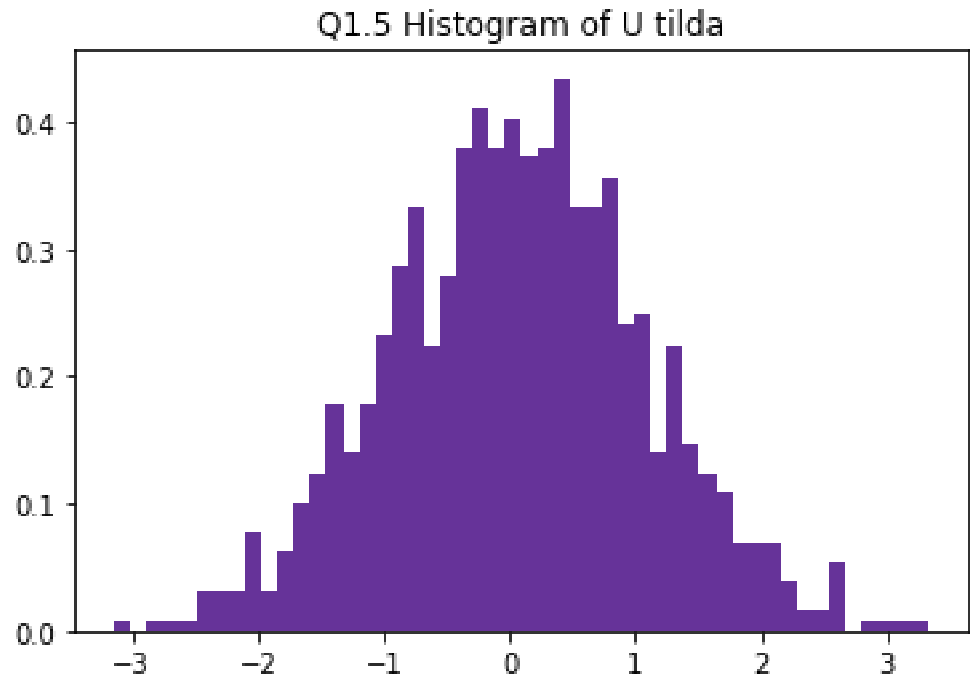
Q1.3: (Note: I ran with the tolerance = 1e-10 instead of 1e-5 to get a smooth graph throughout the whole range)



Q1.4: Yes, the histogram is very similar to Q1.1.



Q1.5: Ui should follow a standard Gaussian distribution approximately.



Q2.1:

Q2.1

Let $q_i = \exp(x^T x_i/\lambda) / \sum_{j=1}^{n} \exp(x^T x_j/\lambda)$

So, $\log q_i = x^T x_i/\lambda - \log\left(\sum_{j=1}^{n} \exp(x^T x_j/\lambda)\right)$

- adding & subtracting $\log q_i$ in the second term

we get

$$\underset{p \in \mathbb{R}^n_+, \sum_i p_i = 1}{\operatorname{argmin}} \underbrace{\sum_{i=1}^{n} p_i \cdot -x^T x_i + \lambda p_i \left(\log p_i - \log q_i + \log q_i\right)}$$

$$= \sum_i p_i \cdot -x^T x_i + \lambda p_i \log \frac{p_i}{q_i} + \lambda p_i \log q_i$$

$$= \sum_i p_i \left(-x^T x_i + \lambda \log q_i\right) + \lambda p_i \log \frac{p_i}{q_i}$$

$$= \sum_i p_i \left(-x^T x_i + x^T x_i - \lambda \log\left(\sum_{j=1}^{n} \exp(x^T x_j/\lambda)\right)\right) + \lambda p_i \log \frac{p_i}{q_i}$$

$$= \sum_i p_i \left(-\lambda \log \sum_{j=1}^{n} \exp(x^T x_j/\lambda)\right) + \lambda p_i \log \frac{p_i}{q_i}$$

Let $-\lambda \log \sum_{j=1}^{n} \exp(x^T x_j/\lambda) = C$ which is independent of i

So the summation becomes

$$C \sum_{i=1}^{n} p_i + \lambda \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i}$$

- $\sum_{i=1}^{n} p_i = 1$ by the constraint & $\sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} = KL(p \| q)$

So the problem becomes

$$\underset{p \in \mathbb{R}^n_+, \sum_i p_i = 1}{\operatorname{argmin}} \; C + \lambda KL(p \| q)$$

which is minimized when $p = q$

since $KL(p \| q) \geq 0$ & achieves equality iff $p = q$

$$\therefore \; p_i = q_i = \frac{\exp(x^T x_i/\lambda)}{\sum_{j=1}^{n} \exp(x^T x_j/\lambda)}$$

Q2.2:

Q 2.2

$$\underset{z \in R^d}{\operatorname{argmin}} \sum_{i=1}^{n} p_i \|z - x_i\|_2^2$$

↳ differentiating w.r.t. z we get

Since

$$\|z - x_i\|_2^2 = \langle z - x_i, z - x_i \rangle$$

$$= \langle z, z \rangle - 2\langle z, x_i \rangle + \langle x_i, x_i \rangle$$

$$\frac{\partial}{\partial z} = \sum_{i=1}^{n} 2p_i \cdot (z - x_i) = \vec{0}$$

$$\rightarrow \sum_{i=1}^{n} 2p_i z = \sum_{i=1}^{n} 2p_i x_i$$

$$\rightarrow z \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} p_i x_i$$

since $\sum_{i=1}^{n} p_i = 1$

$$\rightarrow z = \sum_{i=1}^{n} p_i x_i$$

Q2.3: They do all converge to one point. It seems like the point of convergence is some kind of mean but I'm not sure which.