Sapienza University of Rome

Big Data Computing

Homework 4

Student

Clara Lecce, 1796575

Academic Year 2023/2024

Assignment 1

The all assignment was done on Colab in the file `1796575-Lecce_HW4_2023.ipynb`.

Assignment 2

1)

To prove that $\mathbb{E}\left[f(\mathbf{x})^T f(\mathbf{y})\right] = \mathbf{x}^T\mathbf{y}$, we can proceed by unwrapping the formula:

$$\begin{aligned}
\mathbb{E}\left[f(\mathbf{x})^T f(\mathbf{y})\right] &= \mathbb{E}\left[(S\mathbf{x})^T (S\mathbf{y})\right] \\
&= \mathbb{E}\left[\frac{1}{\sqrt{k}}\mathbf{x}^T U^T \frac{1}{\sqrt{k}} U\mathbf{y}\right] \\
&= \frac{1}{k}\mathbf{x}^T\mathbf{y}\mathbb{E}\left[U^T U\right]
\end{aligned}$$

Let's focus on $\mathbb{E}\left[U^T U\right]$; we can rewrite it as following:

$$\mathbb{E}\left[U^T U\right] = \mathbb{E}\left[\sum_{i=1}^{k} \mathbf{u}_{ij}^T \mathbf{u}_{ij}\right] = \sum_{i=1}^{k} \mathbb{E}\left[\mathbf{u}_{ij}^T \mathbf{u}_{ij}\right]$$

Since $U_{ij} \sim \mathcal{N}(0,1)$ and are independent, with $j = 1, \ldots, d$:

$$\mathbb{E}\left[\mathbf{u}_{ij}^T \mathbf{u}_{ij}\right] = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \quad\quad i \neq j \end{cases}$$

So $\sum_{i=1}^{k} \mathbb{E}\left[\mathbf{u}_{ij}^2\right] = k$, and to continue the unwrapped formula:

$$\begin{aligned}
\mathbb{E}\left[f(\mathbf{x})^T f(\mathbf{y})\right] &= \frac{1}{k}\mathbf{x}^T\mathbf{y}\mathbb{E}\left[U^T U\right] \\
&= \frac{1}{k}\mathbf{x}^T\mathbf{y}k \\
&= \mathbf{x}^T\mathbf{y}
\end{aligned}$$

2)

The cosine of the angle between $\mathbf{x}$ and $\mathbf{y}$ is:

$$cos(\theta) = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$$

We can remove the denominator, since $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$:

$$cos(\theta) = \mathbf{x}^T\mathbf{y}$$

But this is equal to the expectation calculated in the first point, so we have our unbiased estimator with $f(\mathbf{x})$ and $f(\mathbf{y})$, with the expectation equal to:

$$\mathbb{E}\left[f(\mathbf{x})^T f(\mathbf{y})\right] = \mathbf{x}^T\mathbf{y} = cos(\theta)$$