csci-5454-hw6

Jake Krol

October 2024

1 Question 1

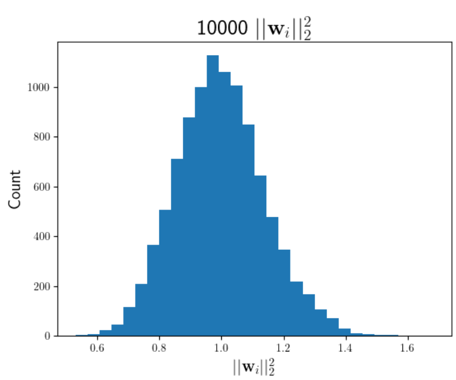
1.1 Question 1a

A fixed random vector, $\mathbf{x}$, was generated by taking 1000 samples of a standard normal distribution, then the vector was divided by its L2 norm to set $||\mathbf{x}||_2 = 1$.

$$\sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_i^2 + \ldots + x_d^2} = ||\mathbf{x}||_2$$

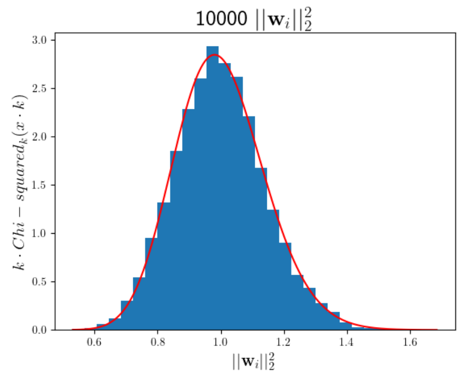
$N = 10000$ Johnson-Lindenstrauss (JL) matrices were generated and multiplied by $\mathbf{x}$ to produce N JL-transformed vectors $\mathbf{w}_i + \ldots \mathbf{w}_N$. The squared L2 norm $||\mathbf{w}_i||_2^2$, was calculated for N transformations.

$$\mathbf{w}_i = G_i\mathbf{x}$$
$$||\mathbf{w}_i||_2^2 = \mathbf{w}_i \cdot \mathbf{w}_i$$



1.2 Question 1b

Each element of a JL-transformed vector is a sum of i.i.d. scaled Gaussian random variables $w_i = \sum_{j=1}^{d} x_j \cdot g_j$ (where w_i is an element of JL-transformed vector, $\mathbf{w}$), and the squared L2 norm involves summing squared Gaussians which results in a Chi-squared random variable.



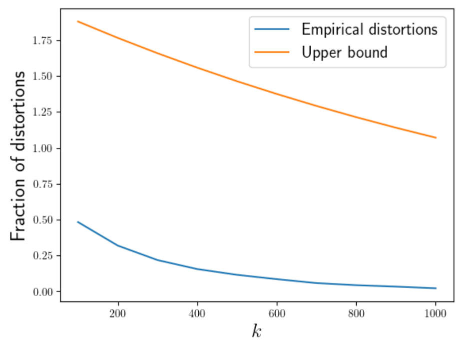
Since $E[g_j] = 0$ and $VAR[g_j] = 1$ the expected value is $E[w_i] = \sum_{j=1}^{d} x_j E[g_j] = 0$ and, given independence, the variance is $VAR[w_i] = E[w_i^2] - E[w_i]^2 = E[w_i^2] - 0 = \sum_{j=1}^{d} VAR[x_j \cdot g_j] = \sum_{j=1}^{d} x_j^2 \cdot VAR[g_j] = \sum_{j=1}^{d} x_j^2 \cdot 1 = x_1^2 + \ldots + x_d^2 = ||x||_2^2$. By linearity of expectation, the expected value of the sum of squared Gaussians is also $||x||_2^2$.

$$||\frac{1}{\sqrt{k}} Gx||_2^2 = \frac{1}{k} \sum_{i=1}^{k} w_i^2$$

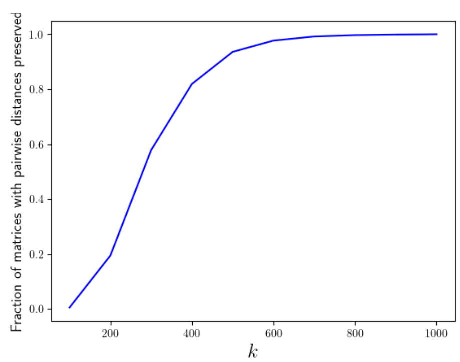
$$E[\frac{1}{k} \sum_{i=1}^{k} w_i^2] = E[\frac{1}{k} \sum_{i=1}^{k} ||x||_2^2] = ||x||_2^2$$

1.2.1 Question 1c

The number of distortions from the empirical simulation and the upper bound $2e^{\frac{-k\epsilon^2}{4}}$ for various k shows the empirical count of distortions is much less than the upper bound. The fraction of distortions represents the probability of $||\mathbf{w}||_2^2$ deviating greater than $\epsilon = 0.05$ from the original norm $||\mathbf{x}||_2^2$, and the results demonstrate well that this probability will not exceed the upper bound.



1.2.2 Question 1d



1.3 Question 2

1.3.1 Question 2a

The algorithm returns HEADS with $\frac{1}{2^k}$ Pr by leveraging k independent Bernoulli random variables. By using k independent Bernoulli random variables $X_i \dots X_k$ each with $p = \frac{1}{2}$ the intersection of all outcomes being HEADS is $p^k = \frac{1}{2^k}$. In contrast, the complementary event is handled by returning TAILS immediately if any of the k tosses are not HEADS.

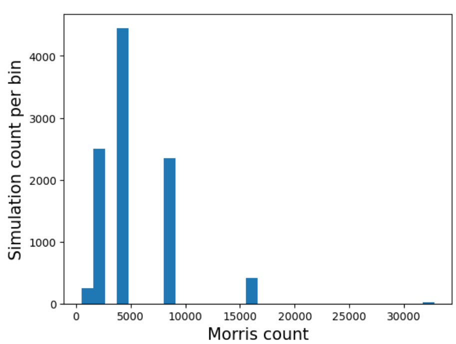
Algorithm 1 BIASED-COIN(k)

```
i = 0
while i < k do
    c = FLIP(p = 1/2)                ▷ flip fair coin
    if c == HEADS then
        i = i + 1
    else
        return TAILS
    end if
end while
return HEADS
```

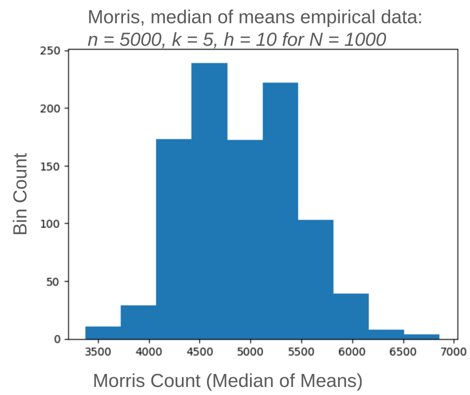
1.3.2 Question 2b

	Theory	$h = 10^4$ simulations
Expected (Avg) Estimated Count	5000	≈ 5032
Standard Deviation	$\sqrt{\frac{n(n-1)}{2}} \approx 3535$	≈ 3583

Table 1: Morris counter empirical results and theory



1.3.3 Question 2c



	Theory	$h = 10(corrected), k = 5$
Expected (Avg) Estimated Count	5000	≈ 4883
Standard Deviation	- do not bother-	≈ 552

Table 2: Morris counter median of means

The variance of the median of means Morris counter (part 2c) is less than the variance for the 10^4 individual Morris counters (part 2b). Particularly, using samples means, $k = 5$ super counters, results in lower variance for each trial. Taking the median super-counter from each trial also aids in centering the final outcome since medians are rank-based which can mitigate the undesired weight toward high magnitude outliers.