

STAT 672: Homework #1

Tom Wallace

February 9, 2018

Problem 1

A

See attached code `hw1_A.py` for an implementation of rejection sampling from the unit cube.

The acceptance probability is the volume of an n -dimensional ball divided by the volume of an n -dimensional cube:

$$P_{X \sim B_\infty^d}(\|X\|_2 \leq 1) = \frac{\text{Vol}(B_2^d)}{\text{Vol}(B_\infty^d)} = \frac{\frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}}{2^d} \quad (1)$$

We know that $\Gamma(n+1) = n!$. Thus, $\Gamma(\frac{d}{2}+1) = (\frac{d}{2})!$

Stirling's approximation for factorials is useful here.

$$n! \approx \sqrt{2\pi n} n^n e^{-n} \quad (2)$$

Let us consider the case of $2d$. In this case:

$$\text{Vol}(B_2^{2d}) = \frac{\pi^d}{\Gamma(d+1)} = \frac{\pi^d}{d!} \approx \frac{\pi^d}{\sqrt{2\pi d} d^d e^{-d}} \quad (3)$$

Rearranging terms, we have:

$$= \frac{1}{\sqrt{2\pi d}} \left(\frac{\pi e}{d}\right)^d \quad (4)$$

Referring to $\left(\frac{\pi e}{d}\right)^d$, the denominator grows faster with d than does the numerator. We can show this more formally by computing the derivative of the numerator and of the denominator, respectively:

$$\frac{d}{dx}((\pi e)^x) = (\pi e)^x (1 + \ln \pi) \quad (5)$$

$$\frac{d}{dx}(x^x) = x^x (1 + \ln x) \quad (6)$$

As x grows larger, (6) becomes increasingly larger than (5). As a consequence, if we extend d infinitely, the limit of the ratio is 0.

$$\lim_{d \rightarrow \infty} \left(\frac{\pi e}{d}\right)^d = 0 \quad (7)$$

Which implies:

$$\lim_{d \rightarrow \infty} \frac{1}{\sqrt{2\pi d}} \left(\frac{\pi e}{d}\right)^d = 0 \quad (8)$$

In other words, as $d \rightarrow \infty$, more and more of the volume of the unit ball is contained near its surface. This phenomenon is called *concentration of measure*.

Returning to (1), we compute the the volume of the unit cube as $d \rightarrow \infty$.

$$\lim_{d \rightarrow \infty} 2^{2d} = \infty \quad (9)$$

So, the ratio of the unit sphere to the unit cube is:

$$\lim_{d \rightarrow \infty} \frac{\text{Vol}(B_2^d)}{\text{Vol}(B_\infty^d)} = \frac{0}{\infty} = 0 \quad (10)$$

Since this ratio also is the probability of accepting a sample in a rejection sampling scheme, our expected runtime to generate a single random vector from B_2^d with rejection sampling grows asymptotically with d . For example, for $d = 10000$, our program would run for a very, very, *very* long time.

B

We can generate Θ as $\Theta = Z/\|Z\|_2$, where $Z \sim N(0, I_d)$ due to the Box-Muller (B-M) transformation.

Problem 2

A

Below are the results of my simulation, the code of which is attached as `hw1_B.py`. As d grows larger, the frequency of the event $\{\|X_+ - Z\|_2 \geq \|X_- - Z\|_2\}$ increases. This is somewhat counter-intuitive. Because Z is drawn from the same distribution as X_+ , we would expect their difference $X_+ - Z$ to result in something close to a zero-vector, which should have a smaller Euclidean norm than that of $X_- - Z$ (since these two have different means), resulting in *low* frequency of the event. Yet, as dimensionality grows, this is not the case: the event occurs with *high* frequency. This has three explanations. One, because the mean is $\pm 5/\sqrt{d}$, as d increases, the means increasingly converge towards 0. Two, X_+ and Z have larger variance than X_- , and this variance does not decrease with d . Thus, as d increases, the larger variance of X_+ and Z predominates over their increasingly similar means, resulting in them more often being further apart (their difference has a larger Euclidean norm) than X_- and Z . Three, the curse of dimensionality means the probability that any two vectors are far apart increases with dimensionality. This magnifies the effect of variance, which as previously discussed is greater for X_+ and Z .

