

# Math 335 HW5

Rohan Mukherjee

November 14, 2023

1. *High dimensional roulette.* I would not play this game. We can show this by showing that the average value of the left hand side is greater than the average value of the right hand side, which would mean that on average, I would lose. The average value of the LHS is

$$\begin{aligned}\int_{[0,1]^{16}} \|x\|^2 dx &= \int_0^1 \cdots \int_0^1 x_1^2 + \cdots x_{16}^2 dx_1 \cdots dx_{16} \\ &= \int_0^1 \cdots \int_0^1 x_1^2 dx_1 \cdots dx_{16} + \cdots + \int_0^1 \cdots \int_0^1 x_{16}^2 dx_1 \cdots dx_{16} \\ &= \frac{16}{3}\end{aligned}$$

Also note that the other 14 integrals have vanished because the inside does not depend on  $x_{17}, \dots, x_{30}$ , and we are integrating over  $[0,1]$ . The result holds as each integral is  $1/3$ . Doing the exact same argument on the RHS, but noting that we only have 14 variables this time (with 16 integrals disappearing), we would get that the average value of the RHS is  $\frac{14}{3}$ , which is indeed less than  $\frac{16}{3}$ . So on average, the LHS is greater than the RHS, which would mean that I lose. As I lose on average, I would not want to play this game. Next, this is the code that I used to simulate

*Low dimensional roulette.* The results on this one were fairly confusing. It seems that in the long run, if you are losing then you start losing a LOT. If you are winning, you start winning a LOT. For example, when I did 1000 runs of 10,000 rounds, and averaged them, sometimes I got 600, and other times I got -600. Therefore, I would probably play, and if I am consistently losing money, I should probably stop, and if I consistently winning money, I should probably continue. I will say this seems awfully intuitive. Here is my code:

```
from numpy import random
import numpy as np

moneys = np.empty([100], dtype=int)
money = 100
for k in range(100):
    for n in range(1000):
        money -= 1
        m = random.randint(1, 3)
        if m == 2:
```

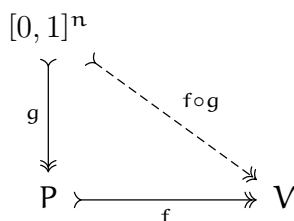
```

    money += 2
    moneys[k]=money
    print(np.average(moneys))

```

I also tried seeing if making it say  $m=2$  vs  $m=1$  made a difference, but it doesn't. It seems like this gambling can be very profitable, but if you start losing, you should definitely stop. In the real world it seems you will keep losing.

2. Given a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  where  $A \in \mathbb{R}^{n \times n}$ , I claim that  $Df = A$  (the constant matrix). We consider the first row of the matrix. The first entry of the matrix product  $Ax$  would be  $\langle a_1, x \rangle$  (where  $a_1$  is the first row of  $A$ ). We have seen in Math 334 that the gradient of this function would be simply  $a_1$ , so the first row of the determinant matrix would be  $a_1$ . Continuing this process, we see that the  $i$ -th row has the constant vector  $a_i$  in it, that is that  $Df(x) = A$  (where the right side doesn't depend on  $x$ ). Next, given any  $A, B \in \mathbb{R}^{n \times n}$ , let  $f(x) = Ax$ , and  $g(x) = Bx$ . Note that  $(f \circ g)(x) = ABx$ , and therefore  $D(f \circ g)(x) = AB$  by our discussion above. Suppose first that  $\det(A) = 0$ . If  $f \circ g$  was onto, then  $f$  is necessarily onto, which would contradict the statement that  $\det(A) = 0$ . So it must be that  $\det(AB) = 0$ . If  $\det(B) = 0$ , then  $B$  is not one-to-one, so there certainly exist vectors  $a \neq b \in \mathbb{R}^n$  so that  $Ba = Bb$ . Then clearly  $f(g(a)) = f(g(b))$ , as the inputs to  $f$  are equal, but as the inputs to  $f \circ g$  are not equal,  $f \circ g$  is not one-to-one, and  $\det(AB) = 0$ . So now suppose that neither are 0. Then  $g$  maps  $[0, 1]^n$  bijectively to some parallelepiped  $P$ , and  $f$  maps  $P$  bijectively to some other parallelepiped  $V$ . Then  $f \circ g$  maps  $[0, 1]^n$  bijectively to the parallelepiped  $V$  (the composition of two bijections is bijective—the two-sided inverse of  $f \circ g$  would be  $g^{-1} \circ f^{-1}$ ). Here is a picture of what's going on:



From this we can say that:

$$\int_V 1 dx = \int_{[0,1]^n} |\det(AB)| dx$$

By our discussion above about  $D(f \circ g)$ , and change of variables. Note also that we could do these one at a time, that is,

$$\int_V 1 dx = \int_P |\det(A)| dx = |\det(A)| \int_P dx = |\det(A)| \int_{[0,1]^n} |\det(B)| dx$$

By simply doing the change of variables twice. Equating the first integral with the second, pulling out the constant and noting that  $[0, 1]^n$  has volume 1, we see that  $|\det(A)| |\det(B)| = |\det(AB)|$ . I tried originally proving that if  $P$  is a parallelepiped of volume  $Q$ , then the

image of  $P$  under  $f$  (that is,  $AP$ ) would have volume  $\det(A) \cdot \text{Vol}(P)$ , which would make a lot of intuitive sense. I believe that is what this formula is saying, but this statement was harder to prove (really just harder to formalize, like what does one mean by "Let  $P$  be a parallelepiped"?), so I went with this instead.

3. Given any list of vectors  $\{v_1, \dots, v_m\}$  so that they are pairwise orthogonal, suppose that there were constants (not all zero)  $a_1, \dots, a_m \in \mathbb{R}$  so that  $a_1 v_1 + a_2 v_2 + \dots + a_m v_m = 0$ . WLOG  $a_1$  is not zero (renumber the vectors if this is not the case). Then we can rearrange to find that  $v_1 = -(a_2/a_1 v_2 + \dots + a_m/a_1 v_m)$ . Now notice that  $1 = \langle v_1, v_1 \rangle = -\langle v_1, a_2/a_1 v_2 + \dots + a_m/a_1 v_m \rangle = -a_2/a_1 \langle v_1, v_2 \rangle - a_3/a_1 \langle v_1, v_3 \rangle - \dots - a_m/a_1 \langle v_1, v_m \rangle$ , and as  $\langle v_k, v_l \rangle = 0$  for all  $k \neq l$ , we see that the RHS of the above expression equals 0, a contradiction. So our assumption that the set was linearly dependent was incorrect, and the set must be linearly independent. We prove that if  $S \subseteq \mathbb{R}^d$ , then  $\dim(S) \leq d$ . We can find a basis for  $S$ , say  $\{s_1, \dots, s_m\}$ . Extend this to a basis for  $\mathbb{R}^d$ , say  $\{s_1, \dots, s_m, q_1, \dots, q_k\}$ . We see that  $m+k = n$ , and as  $k \geq 0$ , we see that  $m \leq n$ . As our vectors above form a basis for its span, and its span has dimension  $m$  by the big theorem of linear algebra, we have proven our statement. Given  $x \in \text{span}\{v_1, \dots, v_m\}$ , we see that  $x = a_1 v_1 + \dots + a_m v_m$  for some scalars  $a_1, \dots, a_m \in \mathbb{R}$ . It suffices to find each  $a_i$ . We see that  $\langle x, v_1 \rangle = \langle a_1 v_1 + \dots + a_m v_m, v_1 \rangle = a_1 \langle v_1, v_1 \rangle + a_2 \langle v_1, v_2 \rangle + \dots + a_m \langle v_1, v_m \rangle$ . All inner products that have different  $v_i$ 's are automatically zero, and as  $\langle v_1, v_1 \rangle = 1$ , we see that  $a_1 = \langle x, v_1 \rangle$ . The same argument works for any  $v_i$ , so we see that in general  $a_i = \langle x, v_i \rangle$ , and we are done.

4. Note that we wish to calculate the value of the integral

$$\int_{B(0,1)} \int_{B(0,1)} (\|X\| - \|Y\|)^2 dX dY / \text{Vol}(B(0,1))^2$$

We proceed by evaluating the inner integral using the radial trick below. Fix  $Y$ , then  $\|Y\| = c$  for some constant  $c$ . We can now say that  $f(X) = (\|X\| - c)^2$ , and therefore  $g(r) = (r - c)^2$ , clearly now  $f(X) = g(\|X\|)$ . Therefore the inner integral above is equal to

$$\begin{aligned} c_d \int_0^1 g(r) \cdot r^{d-1} dr &= c_d \int_0^1 (r - c)^2 \cdot r^{d-1} dr \\ &= c_d \int_0^1 r^{d+1} - 2r^d c + c^2 r^{d-1} dr \\ &= c_d \left( \frac{1}{d+2} - \frac{2c}{d+1} + \frac{c^2}{d} \right) \end{aligned}$$

Now recall that  $c = \|Y\|$ , so the top integral has now become (without the volume scaling)

$$c_d \int_{B(0,1)} \frac{1}{d+2} - \frac{2\|Y\|}{d+1} + \frac{c^2}{d} dY = c_d \left( \frac{\text{Vol}(B(0,1))}{d+2} - \frac{2}{d+1} \int_{B(0,1)} \|Y\| dY + \frac{1}{d} \int_{B(0,1)} \|Y\|^2 dY \right)$$

Now we proceed by exactly the same trick. The first integral:

$$\begin{aligned} \int_{B(0,1)} \|Y\| dY &= c_d \int_0^1 r^d dY \\ &= \frac{c_d}{d+1} \end{aligned}$$

Similarly, the second integral becomes  $\frac{c_d}{d+2}$ . So our expression is now:

$$c_d \left( \frac{\text{Vol}(B(0, 1))}{d+2} - \frac{2c_d}{(d+1)^2} + \frac{c_d}{d(d+2)} \right)$$

Now we simply have to express  $\text{Vol}(B(0, 1))$  in terms of  $c_d$ , and we are done.

$$\begin{aligned} \text{Vol}(B(0, 1)) &= \int_{B(0,1)} 1 dx \\ &= \int_0^1 c_d r^{d-1} dr \\ &= \frac{c_d}{d} \end{aligned}$$

Where we have done the radial transformation on the function  $g(r) \equiv 1$ . Therefore, the above integral above simplifies to

$$c_d \left( \frac{c_d}{d(d+2)} - \frac{2c_d}{(d+1)^2} + \frac{c_d}{d(d+2)} \right)$$

Finally, dividing by the volume squared nets us that the final integral equals

$$d^2 \left( \frac{1}{d(d+2)} - \frac{2}{(d+1)^2} + \frac{1}{d(d+2)} \right) = 2d^2 \left( \frac{1}{d(d+2)} - \frac{1}{(d+1)^2} \right)$$

Which is pretty nice. This can be interpreted as the average value of the difference in magnitudes of two points in the  $d$ -dimensional sphere squared.

5. By what we showed in class, we have that

$$\int_{\mathbb{R}^d} e^{-x_1^2 - x_2^2 - \dots - x_d^2} dx = c_d \int_0^\infty e^{-r^2} r^{d-1} dr$$

The LHS is equal to

$$\begin{aligned} \int_{-\infty}^\infty \int_{-\infty}^\infty \dots \int_{-\infty}^\infty e^{-x_1^2} e^{-x_2^2} \dots e^{-x_d^2} dx_1 \dots dx_d &= \int_{-\infty}^\infty e^{-x_1^2} dx_1 \int_{-\infty}^\infty e^{-x_2^2} dx_2 \dots \int_{-\infty}^\infty e^{-x_d^2} dx_d \\ &= \underbrace{\sqrt{\pi} \dots \sqrt{\pi}}_{d \text{ times}} \\ &= \pi^{\frac{d}{2}} \end{aligned}$$

Now we apply the change of variables  $u = r^2$  to the right side, and by using the  $1 \times 1$  jacobian, we see that the RHS is equal to

$$\frac{c_d}{2} \int_0^\infty e^{-u} u^{\frac{d}{2}-1} du = \frac{c_d}{2} \Gamma\left(\frac{d}{2}\right)$$

By rearranging, we may conclude that

$$c_d = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}$$

First we start with the base case that  $n = 0$ .

$$\Gamma(1) = \int_0^{\infty} t^0 e^{-t} dt = - \lim_{x \rightarrow \infty} e^{-t} \Big|_0^x = 1 - \lim_{x \rightarrow \infty} e^{-x} = 1 = 1!$$

Now suppose that for some  $k \geq 0$ ,  $\Gamma(k+1) = k!$ . We proceed using integration by parts.

$$\begin{aligned} \Gamma(k+2) &= \int_0^{\infty} t^{k+1} e^{-t} dt \\ &= \lim_{x \rightarrow \infty} t^{k+1} e^{-t} \Big|_0^x + (k+1) \int_0^{\infty} t^k e^{-t} dt \\ &= (k+1)k! \\ &= (k+1)! \end{aligned}$$

(The limit vanished because as  $t \rightarrow \infty$ ,  $\frac{t^{k+1}}{e^t} \rightarrow 0$ , and when evaluated at 0,  $t^{k+1} = 0$ ) And we are done.  $\square$