Brendan Neal 1001160236 | nealbre1 2018-09-23

CSC420 Homework 1

Question 1: Nearest Neighbours and the Curse of Dimensionality

Question 1a: Expectation and Variance of V

Final Answer (work below): $E(Z) = \frac{1}{6}, V(Z) = \frac{7}{180}$

We know:

a) $Z = (X - Y)^2$
b) $X \sim Uniform(0, 1)$
c) $Y \sim Uniform(0, 1)$
d) X, Y are independent

Using additive property of expectation $E(A + B) = E(A) + E(B)$ to obtain $E(Z)$

$$E(Z) = E((X - Y)^2)$$
$$\Rightarrow E(Z) = E(X^2 + Y^2 - 2XY)$$
$$\Rightarrow E(Z) = E(X^2) + E(Y^2) - 2E(XY)$$

Since X, Y are independent: $E(XY) = E(X)E(Y)$

$$\Rightarrow E(Z) = E(X^2) + E(Y^2) - 2E(X)E(Y)$$

Since X, Y are sampled from the identical distributions, all their moments must be the same

$$\forall n \in \mathbb{R}, E(X^n) = E(Y^n)$$
$$\Rightarrow E(Z) = 2E(X^2) - 2E(X)^2$$
$$\Rightarrow E(Z) = 2(E(X^2) - E(X)^2)$$

Recall properties of uniform distributions: $X \sim Uniform(a, b)$

$$f(x) = \frac{1}{b - a}$$

Thus for $X \sim Uniform(0, 1)$

$$f(x) = 1$$

Creating an equation for X's n-th moment about 0 ($E(X^n)$)

$$E(X^n) = \int_0^1 x^n f(x)dx = \int_0^1 x^n dx = \frac{1}{n+1} x^{n+1}\Big|_0^1 = \frac{1}{n+1}(1 - 0) = \frac{1}{n+1}$$

Using the above $E(X) = \frac{1}{2}, E(X^2) = \frac{1}{3}$

$$E(Z) = 2(E(X^2) - E(X)^2)$$
$$\Rightarrow 2\left(\frac{1}{3} - \left(\frac{1}{2}\right)^2\right) = 2\left(\frac{1}{3} - \frac{1}{4}\right) = \frac{2}{12}$$
$$\Rightarrow E(Z) = \frac{1}{6}$$

Question 1a, continued…

Definition of variance in terms of expectations

$$V(Z) = E(Z^2) - E(Z)^2$$

Using the additive property of expectation to obtain $E(Z^2)$

$$E(Z^2) = E(((X-Y)^2)^2)$$

$$\Rightarrow E(Z^2) = E((X-Y)^4)$$

$$\Rightarrow E(Z^2) = E(X^4 - 4X^3Y + 6X^2Y^2 - 4XY^3 + Y^4)$$

$$\Rightarrow E(Z^2) = E(X^4) - 4E(X^3Y) + 6E(X^2Y^2) - 4E(XY^3) + E(Y^4)$$

Since X, Y are independently sampled: $E(X^nY^m) = E(X^n)E(Y^m)$

$$\Rightarrow E(Z^2) = E(X^4) - 4E(X^3)E(Y) + 6E(X^2)E(Y^2) - 4E(X)E(Y^3) + E(Y^4)$$

Since X, Y are sampled from identical distributions $\forall n \in \mathbb{R}, E(X^n) = E(Y^n)$

$$\Rightarrow E(Z^2) = 2E(X^4) - 8E(X^3)E(X) + 6E(X^2)^2$$

Using the previous moment equation $E(X^n) = \frac{1}{n+1} \Rightarrow E(X^3) = \frac{1}{4}, E(X^4) = \frac{1}{5}$

$$\Rightarrow E(Z^2) = \frac{2}{5} - \left(\frac{8}{4} * \frac{1}{2}\right) + 6\left(\frac{1}{3}^2\right)$$

$$\Rightarrow E(Z^2) = \frac{2}{5} - 1 + \frac{6}{9}$$

$$\Rightarrow E(Z^2) = \frac{1}{15}$$

Using all prior results to compute $V(Z)$

$$\left(E(Z^2) = \frac{1}{15}\right) \wedge \left(E(Z) = \frac{1}{6}\right) \wedge (V(Z) = E(Z^2) - E(Z)^2)$$

$$\Rightarrow V(Z) = \frac{1}{15} - \frac{1}{36}$$

$$\Rightarrow V(Z) = \frac{7}{180}$$

Brendan Neal 1001160236 | nealbre1 2018-09-23

Question 1b: Expectation and Variance of R

Final Answer (work below): $E(R_d) = \frac{d}{6}$

Let:

$$R_d = \sum_{i=1}^{d} Z_i$$

Expectation of R_d (simplify using the additive property of expectation)

$$E(R_d) = E\left(\sum_{i=1}^{d} Z_i\right) = \sum_{i=1}^{d} E(Z_i)$$

Since all X_i, Y_i are drawn from identical distributions, and all Z_i have the same relationship with X_i and Y_i, all Z_i must also be drawn from identical distributions. Therefore, all Z_i have identical expectations:

$$\forall (i,j) \in \{1 \dots d\} \times \{1 \dots d\} : E(Z_i) = E(Z_j)$$

According to the above, we can turn the summation in $E(R_d)$ into a multiplication of $E(Z)$

$$\big(E(Z) = E(Z_1) = \cdots = E(Z_d)\big)$$

$$\Rightarrow E(R_d) = \sum_{i=1}^{d} E(Z_i) = \sum_{i=1}^{d} E(Z)$$

$$\Rightarrow E(R_d) = dE(Z)$$

Since Z is distributed the same as in part (a)

$$\Rightarrow E(R_d) = \frac{d}{6}$$

Question 1b, continued…

Variance of R_d

$$V(R_d) = E(R_d^2) + E(R_d)^2$$

Computing second moment about zero of R_d

$$E(R_d^2) = E\left(\left(\sum_{i=1}^{d} Z_i\right)^2\right) = E\left(\sum_{i=1}^{d} Z_i^2 + \sum_{i \neq j} Z_i Z_j\right)$$

$$= \sum_{i=1}^{d} E(Z_i^2) + \sum_{i \neq j} E\left(Z_i Z_j\right)$$

$$= \sum_{i=1}^{d} E(Z_i^2) + \sum_{i \neq j} E(Z_i)E\left(Z_j\right)$$

$$= \sum_{i=1}^{d} E(Z^2) + \sum_{i \neq j} E(Z)E(Z)$$

$$= dE(Z^2) + \sum_{i \neq j} E(Z)^2$$

$$= dE(Z^2) + d(d-1)E(Z)^2$$

$$= \frac{d}{15} + \frac{(d^2 - d)}{6}$$

$$= \frac{d}{15} + \frac{(d^2 - d)}{6}$$

$$E(R_d^2) = \frac{5d^2 - 3d}{30}$$

$$V(R_d) = E(R_d^2) + E(R_d)^2 = \frac{5d^2 - 3d}{30} - \frac{d^2}{36}$$

$$\frac{25d^2 - 18d}{180}$$

Question 2: Decision Trees