CSC411 HW1

motjumi

May 2019

1 Nearest Neighbours and the Curse of Dimensionality

(a)

$$\mathbb{E}[Z] = \mathbb{E}[(X-Y)^2] = \mathbb{E}[X^2 - 2XY + Y^2] = \mathbb{E}[X^2] - 2\,\mathbb{E}[XY] + \mathbb{E}[Y^2]$$
$$(\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y], \text{ since X and Y are independent})$$
$$= \frac{1}{3} - \frac{1}{2} + \frac{1}{3} = \frac{1}{6}$$

$$\mathrm{Var}[Z] = \mathbb{E}[Z^2] - \mathbb{E}^2[Z]$$
$$\mathbb{E}[Z^2] = E[(X-Y)^4] = \int_0^1 \int_0^1 (x-y)^4 dxdy = \frac{1}{15}$$
$$\mathbb{E}^2[Z] = \frac{1}{36}$$
$$\mathrm{Var}[Z] = \frac{1}{15} - \frac{1}{36} = \frac{7}{180}$$

(b)

$$\mathbb{E}[Z_1 + ... + Z_d] = \mathbb{E}[Z_1] + ... + \mathbb{E}[Z_d] = d\,\mathbb{E}[Z]$$
$$\mathrm{Var}[Z_1 + ... + Z_d] =$$
$$(Z_1, ..., Z_d \text{ are pairwise independent})$$
$$= \mathrm{Var}[Z_1] + ... + \mathrm{Var}[Z_d] = d\,\mathrm{Var}[Z]$$

(c)

max possible distance $= d$

$\mathbb{E}[R] = \dfrac{1}{6}d$

$\mathrm{SD}[R] = \sqrt{\dfrac{7}{180}d}$

Mean squared distance grows linearly with the dimensionality,
and exceeds the length of a side of the hypercube even in d=6.
Standart deviation grows much slower than the mean.
Thus, in high-dimensional space SD is small relative to the mean
and pairwise distances are about the same.