Practica 1 - Ejercicio 8

"Maldicion de la dimension"

Gerardo Machnicki

August 26, 2019

Volumen de una bola contenido cerca de los bordes

[1] Definamos una bola de radio r en dimension d como los puntos tales que
$B_{r,d} = \{x \in \mathbb{R}^d : x^T x \leq r^2\}$
$B_{r,d} = \{r(\frac{x}{r}) \in \mathbb{R}^d : (\frac{x}{r})^T(\frac{x}{r}) \leq 1\} = \{rz \in \mathbb{R}^d : z^T z \leq 1\}$

Por lo tanto se puede pensar en la bola unidad como la bola estrechada en todas las d direcciones en una cantidad r y el volumen de la bola es
$V_{r,d} = r^d V_{1,d}$

La proporción del volumen de una bola unidad que esta contenida en la zona cerca de los bordes se puede pensar como el volumen entre las esferas de radio 1 y de radio $1 - \epsilon$ es $\frac{V_{1,d} - V_{1-\epsilon,d}}{V_{1,d}} = \frac{V_{1,d}[1^d - (1-\epsilon)^d]}{V_{1,d}} = 1 - (1-\epsilon)^d$
La proporción del volumen de una bola unidad en $\mathbb{R}^2$ que esta contenida en una zona de la bola limitante con los bordes y de grosor $\epsilon = 0.05$ sera $(1 - (1 - 0.05)^2) = 0.0975$ mientras que en $\mathbb{R}^{16}$ esta proporción sera $(1 - (1 - 0.05)^{16}) = 0.559$

[1] https://barumpark.com/blog/2018/the-curse-of-dimensionality/

Probabilidad de encontrar un vector normal estandar en la bola unidad $B_1 = \{u \in \mathbb{R}^p : x^T x \leq 1\}$ en distintas dimensiones

Para $p = 2$ y con $x \in \mathbb{R}^2 \sim N_2(0, \mathbb{I})$, $x = (x_1, x_2)$
$Prob(x \in B_1) = Prob(x^T x \leq 1)$
Como $x^T x = x_1^2 + x_2^2$ y como $x_1 \sim N(0,1)$ y $x_2 \sim N(0,1)$ ya que x_1 y x_2 son componentes independientes de una Normal mutivariada con $\Sigma = \mathbb{I}$, entonces estamos sumando normales al cuadrado y tenenos que
$x^T x = x_1^2 + x_2^2 = \chi_2^2$

Para p dimensiones con vector $x \in \mathbb{R}^p$ tenemos $x \sim N_p(0, \mathbb{I}_\shortmid)$,
$x = (x_1, x_2, ..., x_p)$ y $x^T x = x_1^2 + x_2^2 + ... + x_p^2 = \chi_2^p$
Considerando $z = x^T x$, buscamos $P(z \leq r^2)$, o sea en este caso $P(z \leq 1)$ con $z \sim \chi_p^2$
Llamando $W = \sqrt{z}$, tenemos
$Prob[W \leq r] = Prob[Z \leq r^2] = Prob[x^T x \leq r^2]$ y
$K_d(w) = Prob[W \leq w] = Prob[\sqrt{Z} \leq w] = Prob[Z \leq w^2] = G_d(w^2)$, con $G_d(.)$ FDA.

Calculo de la probabilidad encontrar un vector normal estándar en la bola unidad $B_1 = \{u \in \mathbb{R}^p : x^T x \leq 1\}$ en distintas dimensiones

```
df <- data.frame(matrix
(ncol = 2, nrow = 0))
x <- c("GL", "Prob xTx <=1"
colnames(df) <- x
for( i in 1 : 16 ){
df[i,1]=i
df[i,2]=pchisq(1, df = i)}
```

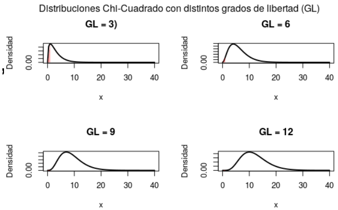

Gráfico de la probabilidad encontrar un vector normal estándar en la bola unidad $B_1 = \{u \in \mathbb{R}^p : x^T x \leq 1\}$ en funcion de la dimension