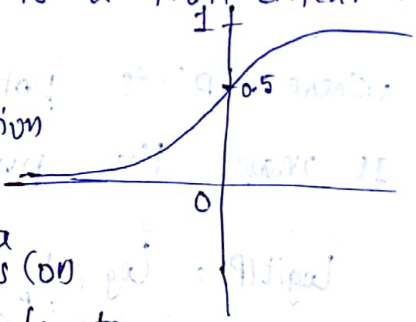
③ The Sigmoid function $\sigma(z) = \frac{1}{1+e^{-z}}$ is a Non-Linear function



In Logistic regression ⇒ It is a classification problem
i.e., classifying our data like giving labels (or) Yes/No, 0/1 etc

The Sigmoid function is S-shaped & bounded function also called as Squashing function - means it maps the whole values into finite interval.

→ so, usually in logistic regression - we use probability values. i.e., $0 \leq p(x) \leq 1$.

This can be acheived by Sigmoid function.

In logistic regression ⇒ we need to give decision boundary to classification of dataset. This can be done by Sigmoid function easily by give $p = 0.5$ as decision boundary.

for any value $p \geq 0.5$ as '1',
$p < 0.5$ classified as '0'.

that means it is used as odd ratios

$$\text{odd}(p) = \frac{p}{1-p} \Leftarrow \frac{\text{probability of success}}{\text{failure probability}}$$

⇒ Very closely related to Binomial distribution.

$$p = \begin{cases} p \geq 0.5 \Rightarrow \text{'1'} \\ p < 0.5 \Rightarrow \text{'0'} \end{cases} \Rightarrow \text{Logistic regression}$$

(4) Logit function means log-odds is the logarithm of odds $\frac{p}{1-p}$

where 'p' is probability.

It maps the probability values from $[0,1]$ to $[-\infty, +\infty]$

finite space — Infinite space

$$\text{Logit}(P) = \log\left(\frac{p}{1-p}\right)$$ → It is like inverse of the Sigmoid 'logistic' function.

$$z = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots$$

⇒ Let $\text{logit}(p) = z$.

$$\Rightarrow \log\left(\frac{p}{1-p}\right) = z = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots$$

write exponent on both sides.

$$\Rightarrow e^{\log_e\left(\frac{p}{1-p}\right)} = e^z$$

$$\Rightarrow \left(\frac{p}{1-p}\right) = e^z$$

⇒ Inverse the equation on bothsides.

$$\frac{1-p}{p} = \frac{1}{e^z}$$

$$\Rightarrow \frac{1-\cancel{p}}{p} = \frac{1}{e^z}$$

$$\Rightarrow \frac{1}{p} = 1 + \frac{1}{e^z} = \frac{1+e^z}{e^z}$$

$$\Rightarrow p = \frac{e^z}{1+e^z}$$ ⇒ divide ~~bothsides~~ numerator & denominator by e^z.

$$\Rightarrow p = \frac{\left(\frac{e^z}{e^z}\right)}{\left(\frac{1}{e^z} + \frac{e^z}{e^z}\right)} = \frac{1}{1+e^{-z}}$$ ← sigmoid function which maps to finite values for logistic regression.

$$\Rightarrow \boxed{p = \frac{1}{1+e^{-z}}}$$ used for Logistic Regression.

5 Entropy of Multivariate Gaussian Variable:-

is a generalization of 1D (univariate) to Higher Dimensions (K-D) (multivariate)

is that a Random-vector is said to be K-variate normally distributed if every linear combination of its 'K' components has a univariate normal distribution.

⇒ Here given dimensionality (k) is 'D'.

Random vector $X = [x_1\ x_2\ \cdots\ x_D]^T$

⇒ $X \sim \mathcal{N}_D(\mu, \Sigma)$

μ = mean of 'D' dimensions

Σ = covariance matrix $(D \times D)$

$= E[(X-\mu)(X-\mu)^T]$

Gaussian distribution is given by.

$$p(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma}\exp\left[\frac{(x-\mu)^2}{2\sigma^2}\right]$$

⇒ with μ = mean

σ = variance.

Here $$p(x) = \frac{1}{\sqrt{2\pi}^D\sqrt{|\Sigma|}}\cdot e^{\frac{-(x-\mu)(x-\mu)^T}{2\Sigma}}$$

⇒ μ = mean

Σ = variance matrix

Entropy $$H(x) = -\int_{-\infty}^{\infty} p(x)\ln(p(x))\,dx.$$

Let $$\ln(p(x)) = \ln\left(\frac{1}{\sqrt{2\pi}^D\sqrt{\Sigma}}\cdot e^{-\frac{(x-\mu)(x-\mu)^T}{2|\Sigma|}}\right)$$

$$\ln(p(x)) = \ln\left(\frac{1}{\sqrt{2\pi}^D\cdot\sqrt{|\Sigma|}}\right) + \frac{(x-\mu)(x-\mu)^T}{2|\Sigma|}$$

$$\Rightarrow H(x) = -\int_{-\infty}^{\infty} p(x)\ln(p(x))\,dx$$

$$= -\int_{-\infty}^{\infty} p(x)\left[\ln\left(\frac{1}{\sqrt{2\pi}^D\cdot\sqrt{|\Sigma|}}\right)\right]dx + \int_{-\infty}^{\infty} p(x)\frac{(x-\mu)(x-\mu)^T}{2|\Sigma|}\,dx.$$

$$= -\int_{-\infty}^{\infty} p(x)\, \ln\left(\frac{1}{\sqrt{2\pi}^{D}\sqrt{|\Sigma|}}\right) dx + \int_{-\infty}^{\infty} \frac{(x-\mu)(x-\mu)^T}{2|\Sigma|}\, p(x)\, dx$$

$$= -\ln\left(\frac{1}{\sqrt{2\pi}^{D}\sqrt{|\Sigma|}}\right) \underbrace{\int_{-\infty}^{\infty} p(x)\, dx}_{=1} + \frac{(x-\mu)(x-\mu)^T}{2|\Sigma|} \underbrace{\int_{-\infty}^{\infty} p(x)\, dx}_{=1}$$

$$= -\ln \frac{1}{\sqrt{2\pi}^{D}\sqrt{|\Sigma|}} + \frac{|\Sigma|}{2|\Sigma|}$$

$$= -\ln \left((\sqrt{2\pi})^{D}\sqrt{|\Sigma|}\right)^{-1} + \frac{1}{2}$$

$$= \ln \left(\sqrt{2\pi}^{D}\sqrt{|\Sigma|}\right) + \frac{1}{2}$$

$$= \ln (2\pi)^{D/2} + \ln |\Sigma|^{1/2} + \frac{D}{2}$$

$$= \frac{D}{2}\ln(2\pi) + \frac{1}{2}\ln(|\Sigma|) + \frac{D}{2}$$

$$= \frac{D}{2}\left[\ln(2\pi)+1\right] + \frac{1}{2}\ln|\Sigma|$$

$$\boxed{H[x] = \frac{1}{2}\ln|\Sigma| + \frac{D}{2}\left[1+\ln(2\pi)\right]}$$

$\Rightarrow$ Entropy of the multivariate Gaussian variable $x \sim \mathcal{N}_D(\mu, \Sigma)$.

Scanned by CamScanner