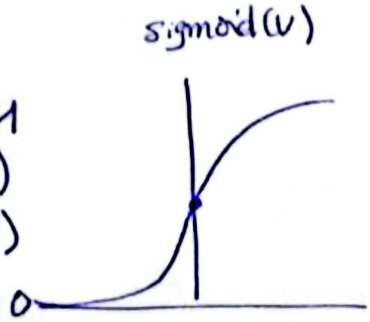
Logistic Regression

- Logistic Regression = classification
- $y = \sigma(a_0 + a_1 x_1 + a_2 x_2 \cdots)$

$\sigma = \dfrac{1}{1+e^{-v}}$, $v = a_0 + a_1 x_1 + a_2 x_2 \cdots$

↓

sigmoid function



if $v > 0 \rightarrow \sigma > 0.5$ (class 1)

if $v < 0 \rightarrow \sigma < 0.5$ (class 0)

↓ probability of an input ↓ binary values

To clarify :

$$\text{Data} \xrightarrow[\text{phase 1: linear Regression}]{\text{2 phases}} y = a_0 + a_1 x_1 \xrightarrow[\text{Logistic Regression}]{} y' = \sigma(y) = \frac{1}{1+e^{-y}}$$

$0 \leq y' \leq 1$

- Loss function : categorical cross-entropy loss function

$$L(y_i, y'_i) = -y_i \log(y'_i) - (1-y_i)\log(1-y'_i)$$

cost function :

$$C(y_i, y'_i) = -\frac{1}{n} \sum y_i \log(y'_i) + (1-y_i)\log(1-y'_i)$$

Concept	Formula
Probability	$p = P(y=1 \mid x)$ $[0, 1]$
Odds	$\frac{p}{1-p}$: prob. of success / prob. of failure $[0, +\infty[$
Logit	$\log\left(\frac{p}{1-p}\right) = z$ "$(\ln(p/1-p))$", $z = \beta_0 + \beta_1 x$ $]-\infty, +\infty[$
sigmoid	$\sigma(z) = 1/1+e^{-z}$
prediction	if $p \geq 0.5$: class 1, else class 0
Loss	$-[y\log(p) + (1-y)\log(1-p)]$

- $\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X$

$$\Leftrightarrow \frac{P}{1-P} = e^{\beta_0+\beta_1 X}$$

$$\Leftrightarrow P = \frac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}} = \frac{1}{1+e^{-(\beta_0+\beta_1 X)}}$$

sigmoid function

Linear Regression Recap

① coefficients

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\Sigma[(x_i-\bar{x})(y_i-\bar{y})]}{\Sigma[(x_i-\bar{x})^2]} = \frac{\Sigma x_i y_i - n\bar{x}\bar{y}}{\Sigma x_i^2 - n\bar{x}^2} = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = \frac{\Sigma y - \beta_1 \Sigma x}{n}$$

② Error variance estimator

$$\sigma^2 = \frac{SSE}{n-2} = \frac{SSE}{df}$$

standard error of estimate

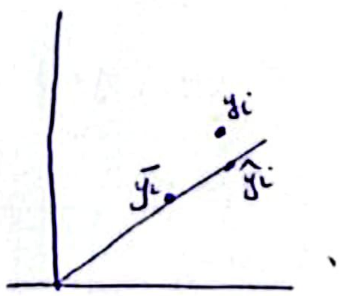
$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{SSE}{n-2}}$$

③ R^2: coefficient of determination

$$R^2 = \frac{SSR}{SST} \quad 0 < R^2 < 1 \quad = 1 - \frac{SSE}{SST}$$

$$SST = SSR + SSE$$

$$\Sigma(y_i-\bar{y})^2 = \Sigma(\hat{y}_i-\bar{y})^2 + \Sigma(y_i-\hat{y}_i)^2$$



SSE ↘, R^2 ↗

④ Hypothesis testing of β_0 & β_1
↳ Standard error of β_0 and β_1
↳ Confidence Interval.