Solving $\hat{\beta}$ is not a linear equation, so you'd need an iterative procedure.

$y_1, \ldots, y_n$ indep Bernoulli (μ_i).

$\boxed{\text{Deviance}}$ – $2\sum_{i=1}^{n}\{y_i \log \frac{y_i}{\mu_i} + (1-y_i)\log\left(\frac{1-y_i}{1-\mu_i}\right)\}$

derivation

0 1

* logit link function for this model.

$= 2\sum_{i=1}^{n} \{-y_i \cdot \log(\mu_i) - (1-y_i)\log(1-\mu_i)\}$

$= 2\sum_{i=1}^{n} \{-y_i \boxed{\log\left(\frac{\mu_i}{1-\mu_i}\right)} - \log(1-\mu_i)\}$

logit of μ_i.

$= 2\sum_{i=1}^{n} \{-y_i \boxed{x_i^T \beta} - \log(1-\mu_i)\}$

the logit function replaces

$= -2\left(\sum_{i=1}^{n} y_i x_i^T\right)\hat{\beta} - 2\sum_{i=1}^{n} \log(1-\hat{\mu}_i)$

$= -2\left(\sum \hat{\mu}_i x_i^T\right)\hat{\beta} - 2\sum_{i=1}^{n} \log(1-\hat{\mu}_i)$

* Notice the y_i's are not present. We don't even need y_i's to get the deviance. Just the known and fixed X design matrix and $\hat{\beta}$. observations. We do need y for $\hat{\beta}$. But after that y is unnecessary.

This deviance is a measure of the fitted values $\hat{\mu}_i$ only.

* [NOT a measure of distance btwn fitted and observed values. CANNOT use it for GOF!