CS6462

Probabilistic and Explainable AI

Lesson 14

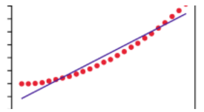
Bayesian Generalized Linear Models

*

Generalized Linear Model Theory

by Emil Vassev

February 24, 2024



LM Generalization

Generalization of the linear regression model:*

- general linear regression model:

$Y = \{y_1, y_2, \dots, y_k\}$ - dependent variable values (observations or realizations)

$X = \{x_1, x_2, \dots, x_n\}$ – independent (explanatory) variable (predictor) and sample set

$Y = N(\mu, \sigma^2)$ $- Y$ has a normal distribution with **mean** μ and variance σ^2

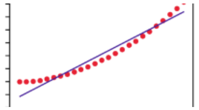
- expected value μ is a linear function of a predictor that takes values $X = \{x_1, x_2, \dots, x_n\}$:

$$\mu = X\beta = \beta_0 + x_i\beta_1$$

$\beta = \{\beta_0, \beta_1\}$ – scale factors (unknown parameters)

- *Extends the linear regression model in two steps:*
    - Generalization based on the Exponential Family of Distributions
    - Generalization based on the link function

J. A. Nelder and R. W. M. Wedderburn (1972). Generalized Linear Models. Journal of the Royal Statistical Society. Series A (General), Vol. 135, No. 3, pp. 370-384



LM Generalization - Exponential Family of Distribution

Mathematical Definition of EFD: formal canonical model

- *meta-exponential model:* different distribution models are generalized
  - one exponential model that unifies Poisson, Normal, Bernoulli, etc.
- random sample $\mathrm{X} = \{\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_\mathrm{n}\}$ with assumed probability distribution $\mathrm{P(X)}$ that depends on some unknown parameter θ
- $f(X_i; \theta)$ – the probability density (or mass) function of each $X_i \in X$ with θ has an exponential form

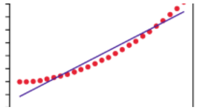
Exponential meta-model for $f(X_i; \theta)$: general exponential forms (canonical forms)

$$f(X_i; \theta) = \exp[\mathrm{a(X_i)} * \mathrm{b}(\theta) + \mathrm{c}(\theta) + \mathrm{d(X_i)}]$$

$$f(X_i; \theta) = e^{\mathrm{a}(x_i) * \mathrm{b}(\theta) + \mathrm{c}(\theta) + \mathrm{d}(x_i)}$$

$\mathrm{a(X_i)}$, $\mathrm{b}(\theta)$, $\mathrm{c}(\theta)$, $\mathrm{d(X_i)}$ - arbitrary functions

LM Generalization – Link Function



LM models the mean:

$$\mu = X\beta = \beta_0 + x_i\beta_1$$

Generalization that shifts the focus from modeling the mean μ :

- introduced one-to-one continuous differentiable transformation $\mathrm{g}(\mu)$

$$\eta = \mathrm{g}(\mu)$$

$\mathrm{g}(\mu)$ - link function (identity, log, inverse, logit, etc.)

$\eta = X\beta$ – linear model

η - linear predictor

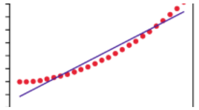
A link function maps a non-linear relationship to a linear one and we can fit a linear model to the data.

- the inverse version of g is used to calculate μ:

$$\mu = g^{-1}(X\beta)$$

Note that we do not transform the observation μ, but rather its expected value μ.

LM Generalization – Canonical Link



Simple generalized linear model:

- relies on the simplest link function - the **identity link function** :

$\eta = g(\mu) = \mu$

- the model -the standard linear model with identity link function:

$\eta = \mu$

Exponential canonical form:

$f(X_i; \theta) = \exp[\mathrm{a}(\mathrm{X_i}) * \mathrm{b}(\theta) + \mathrm{c}(\theta) + \mathrm{d}(\mathrm{X_i})]$

Canonical link:

- case when μ is the same as θ

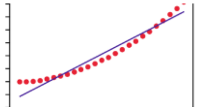
$\mu = \theta$

- the link function makes the linear predictor η the same as the canonical parameter θ

$\mathrm{g}(\mu) = \mu$

$\eta = \mathrm{g}(\mu) = \mu = \theta$

The Generalized Linear Model



Components of GLM:

$F(Y) = \mu = g^{-1}(X\beta)$

μ – expected value of the dependent variable (response variable)

$Y = \{y_1, y_2, \dots, y_k\}$ - dependent variable values (observations)

$X = \{X_1, X_2, \dots, X_n\}, X_i = \{x_1, x_2, \dots, x_k\}$ – independent (explanatory) variables and sample sets

$\beta = \{\beta_0, \beta_1, \beta_2, \dots, \beta_n\}$ – scale factors

- *random component* $F(Y)$ (conditional distribution of the dependable variable)
  - member of the *exponential family of distributions*: e.g., Binary, Binomial, Poisson, Normal, etc.
- *linear predictor* $X\beta$ (a linear function of regressors)

$\eta = X\beta = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 \dots \beta_n * x_n$

- *link function* g -> $g^{-1}(X\beta) = g^{-1}(\beta_0 + \beta_1 * x_1 + \beta_2 * x_2 \dots \beta_n * x_n)$
  - transforms the expectation of the dependent variable *μ ≡ F(Y)* to the linear predictor
  - $g^{-1}(\eta)$ - inverse link function (mean function) maps the linear predictor to the mean

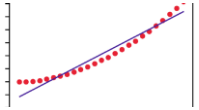
The Generalized Linear Model (cont.)

Common Link Functions and Their Inverses:

- *link function* g -> $g^{-1}(X\beta) = g^{-1}(\beta_0 + \ \beta_1 * x_1 + \beta_2 * x_2 \ .... \ \beta_n * x_n)$
- μ is the expected value of the response; η is the linear predictor; and Φ(·) is the cumulative distribution function of the standard-normal distribution

Link	$\eta_i = g(\mu_i)$	$\mu_i = g^{-1}(\eta_i)$
Identity	μ_i	η_i
Log	$\log_e \mu_i$	e^{η_i}
Inverse	μ_i^{-1}	η_i^{-1}
Inverse-square	μ_i^{-2}	$\eta_i^{-1/2}$
Square-root	$\sqrt{\mu_i}$	η_i^2
Logit	$\log_e \dfrac{\mu_i}{1-\mu_i}$	$\dfrac{1}{1+e^{-\eta_i}}$
Probit	$\Phi^{-1}(\mu_i)$	$\Phi(\eta_i)$
Log-log	$-\log_e[-\log_e(\mu_i)]$	$\exp[-\exp(-\eta_i)]$
Complementary log-log	$\log_e[-\log_e(1-\mu_i)]$	$1-\exp[-\exp(\eta_i)]$

Normal - Identity
Binomial - Logit
Poisson - Log



Summary

Bayesian Generalized Linear Models – Generalized Linear Model Theory

Linear Model Generalization (Nelder and Wedderburn):

- Exponential Family of Distribution
- Link Function
- Canonical Link

Generalized Linear Model:

$$\mathrm{F}(Y) = \mu = g^{-1}(X\beta)$$

Next Lesson:

- Bayesian Nonparametric Models - Gaussian Process Regression Model

Thank You!

Questions?