Multiple Linear Regression

General Principles

To study relationships between multiple continuous independent variables (e.g., the effect of weight and age on height), we can use a multiple regression approach. Essentially, we extend Linear Regression for continuous variable by adding a regression coefficient β_x for each continuous variable (e.g., β_{weight} and β_{age}).

Considerations

> Caution
>
> - We have the same considerations as for the Regression for continuous variable.
> - The model interpretation of the regression coefficients β_x is considered for fixed values of the other independent variable(s)' regression coefficients—i.e., for a given age, β_{weight} represents the expected change in the dependent variable (height) for each one-unit increase in weight, holding all other variable(s) constant (age).

Example

Below is example code demonstrating Bayesian multiple linear regression using the Bayesian Inference (BI) package. Data consist of three continuous variables (*height*, *weight*, *age*), and the goal is to estimate the effect of *weight* and *age* on *height*.

Python

```python
from BI import bi

# Setup device------------------------------------------------
m = bi(platform='cpu')

# Import Data & Data Manipulation ------------------------------------------------
from importlib.resources import files
# Import
data_path = files('BI.resources.data') / 'Howell1.csv'
m.data(data_path, sep=';')
m.df = m.df[m.df.age > 18] # Manipulate
m.scale(['weight', 'age']) # Scale

# Define model ------------------------------------------------
def model(height, weight, age):
    # Parameter prior distributions
    alpha = m.dist.normal(0, 0.5, name = 'alpha')
    beta1 = m.dist.normal(0, 0.5, name = 'beta1')
    beta2 = m.dist.normal(0, 0.5, name = 'beta2')
    sigma = m.dist.uniform(0,50, name = 'sigma')
    # Likelihood
    m.normal(alpha + beta1 * weight + beta2 * age, sigma, obs = height)

# Run MCMC ------------------------------------------------
m.fit(model) # Optimize model parameters through MCMC sampling

# Summary ------------------------------------------------
m.summary()
```

R

```r
library(BI)
m=importbi(platform='cpu')

# Import Data & Data Manipulation ------------------------------------------------
m$data(paste(system.file(package = "BI"),"/data/Howell1.csv", sep = ''), sep=';')# Import
m$df = m$df[m$df$age > 18,] # Manipulate
m$scale(list('weight', 'age')) # Scale
m$data_to_model(list('weight', 'height', 'age')) # Send to model (convert to jax array)

# Define model ------------------------------------------------
```

```r
model <- function(height, weight, age){
  # Parameter prior distributions
  alpha = bi.dist.normal( 0, 0.5, name = 'a')
  beta1 = bi.dist.normal( 0, 0.5, name = 'b1')
  beta2 = bi.dist.normal(  0, 0.5, name = 'b2')
  sigma = bi.dist.uniform(0, 50, name = 's')
  # Likelihood
  m$normal(alpha + beta1 * weight + beta2 * age, sigma, obs=height)
}

# Run MCMC ------------------------------------------------
m$run(model) # Optimize model parameters through MCMC sampling

# Summary ------------------------------------------------
m$summary() # Get posterior distributions
```

Mathematical Details

Frequentist formulation

We model the relationship between the independent variables $(X_{1i}, X_{2i}, ..., X_{ni})$ and the dependent variable Y using the following equation:

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_n X_{ni} + \sigma_i$$

Where:

- Y_i is the dependent variable for observation i.
- α is the intercept term.
- X_{1i}, X_{2i}, ..., X_{ni} are the values of the independent variables for observation i.
- β_1, β_2, ..., β_n are the regression coefficients.
- σ_i is the error term for observation i.

Bayesian formulation

In the Bayesian formulation, we define each parameter with priors . We can express the Bayesian model as follows:

$$Y \sim Normal(\alpha + \sum_{k}^{n} \beta_k X, \sigma^2)$$

$$\alpha \sim Normal(0, 1)$$

$$\beta_k \sim Normal(0, 1)$$

$$\sigma \sim Uniform(0, 50)$$

Where:

- Y_i is the dependent variable for observation i.
- α is the prior distribution for the intercept.
- β_k are the prior distributions for the k distinct regression coefficients.
- X_{1i}, X_{2i}, …, X_{ni} are the values of the independent variables for observation i.
- σ is the prior distribution for the standard deviation, ensuring that it is positive.

Reference(s)

McElreath (2018)

McElreath, Richard. 2018. *Statistical Rethinking: A Bayesian Course with Examples in r and Stan*. Chapman; Hall/CRC.