Logistic regression

Using python.

Iván Andrés Trujilllo Abella

Facultad de Ingenieria
Pontificia Universidad Javeriana

trujilloiv@javeriana.edu.co
addajaveriana

Introduction

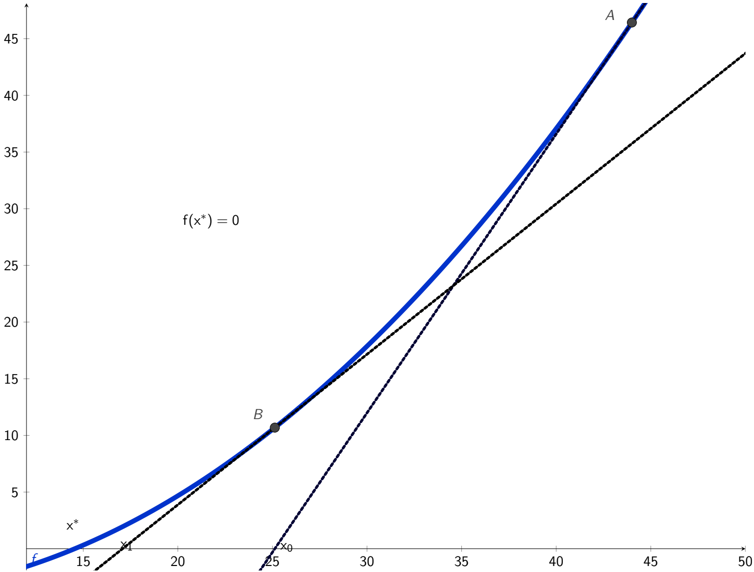
Logistic regression is used broadly in empirical works, it used in economics, engineering, epidemiology and clinical research. In a simplified way the logistic regression it is used to binary problems.

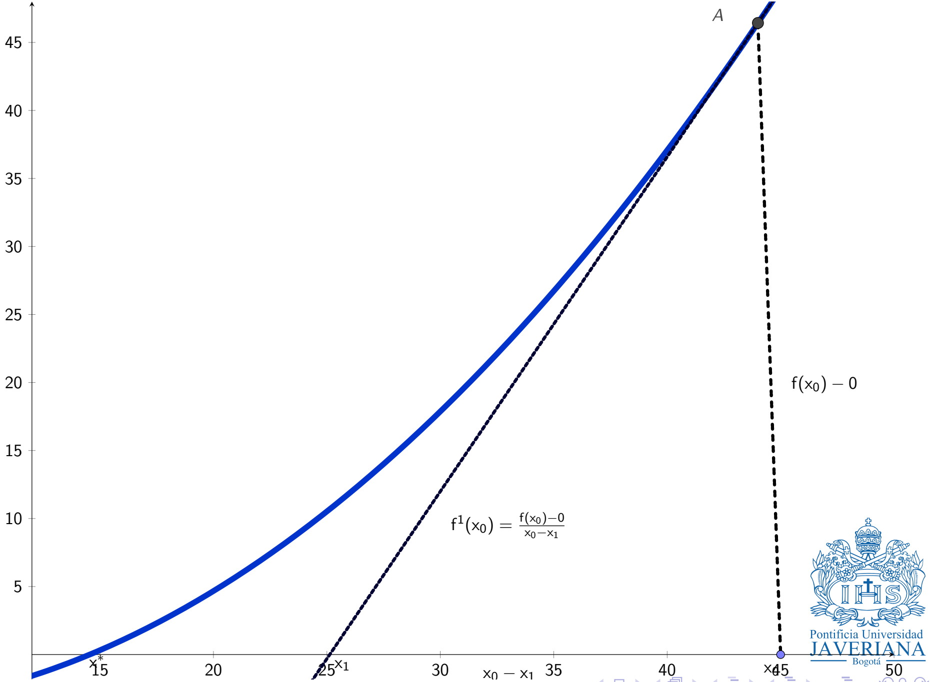
Insights

If we take a point near to the change in concavity the method could produce divergence.

Bolzano theorem

illustration




A
45
40
35
30
25
20
15
10
5
15
20
25
30
35
40
45
50
x^*
x_1
x_0
$x_0 - x_1$
$f(x_0) - 0$
$f^1(x_0) = \frac{f(x_0)-0}{x_0-x_1}$

$$x_1 = x_0 - \frac{f(x_0)}{f^1(x_0)} \tag{1}$$

Thus in i iteration we have:

$$x_{i+1} = x_i - \frac{f(x_i)}{f^1(x_i)} \tag{2}$$

We select the point A the tangent line cross the x-axi in x_0 and again in $f(x_0)$ and again the tangent line of this point (B) cross the x-axis in x_1 each step is also known as a iteration. The algorithm converge to the the value x^* think that this could be used to *optimization problems*.

Newton-raphson program

```
def quadratic(a,b,c, x):
     return a*x**2 + b*x + c
def dfQuadratic(a,b,x):
  return a*x**2 + b*x
a,b,c ,x0 = 1,-3,-4,8
def raphsonQuadratic(a,b,c,x0, error_max=0.0000015,
    iteration_max=100):
  xi = x0
  iter, error = 0, 100
  data = []
  while (iter < iteration_max) and (error > error_max):
    xj = xi - quadratic(a,b,c,xi) / dfQuadratic(a,b,xi)
    error = abs(xj - xi)
    iter += 1
    data.append((xj,xi,error,iter))
    xi = xj
  return data
raphsonQuadratic(a,b,c,4)[-1]
```

second order approximation

taylor series

Insights abour classic and bayesian

Uses likelihood function estimation pretend assume that
$P(A \mid B) = P(B \mid A)$

$P(\Theta)$ prior distribution, posterior distribution $P(\Theta \mid \mathrm{X})$. likelihood $P(\mathrm{X} \mid \Theta)$. Prior the belief before seen the data.

Conjugate prior

Gamma to poison

$x \sim Poisson(\lambda)$

Max a posteriori

A set of features $X = \{x_1, ...x_n\}$ assuming a distribution $P(X, \Theta)$ where *Theta* is a parameter (a random variable).

$$\Theta_{ma} = \max P(\Theta \mid X) \tag{3}$$

the last equation must be compared regarding the maximun likelihood estimation. that establish the $\max P(X \mid \Theta)$.

Pontificia Universidad JAVERIANA Bogotá

Bayesian continuous

$$f_{X|Y}(x \mid y) = \frac{f_{Y|X}(y \mid x) f_X(x)}{f_Y(y)} = \frac{f_{Y|X}(y \mid x) f_X(x)}{\int f_{Y|X}(y \mid x) f_X(x) dx} \tag{4}$$

Maximun Likelihood Estimation

MLE

Suppose that you data it is generated by a theoretical distribution, the inverse problem is determine the most probable parameter that generate the data.

Problem

you have n balls in a bag where there are j reds and k black thus $n = j + k$. However you do not know the really proportion of each one color. if you draw balls and both are different what is the proportion of black balls Θ.

MLE insight

Then we choose a Θ among all posible values that maximize the probability of seen the data.

Problem

In a formal way we can assume that $x \sim B(n, \Theta)$ and we have seen the data results $\{x_1, x_2, x_3, ..., x_n\}$

$$P(X = x_1 \cap X = x_2, ..., \cap X = x_n) \tag{5}$$

The variables are i.i.d and therefore the joint probability is the result of multiply the marginal probabilities.

$$P(X = x_1 \cap X = x_2, ..., \cap X = x_n) = \prod_{i=1}^{n} P(x_i) \tag{6}$$

Problem

$$
\begin{aligned}
\prod_{i=1}^{n} P(x_i) &= \prod_{i=1}^{n} \binom{m}{x_i} \Theta^{x_i} (1-\Theta)^{m-x_i} \\
&= \prod \binom{m}{x_i} \prod \Theta^{x_i} \prod (1-\Theta)^{m-x_i} \\
&= \prod \binom{m}{x_i} \Theta^{\sum x_i} (1-\Theta)^{\sum (m-x_i)} \\
ln(\prod P(x_i)) &= ln(\binom{m}{x_i} + \sum x_i ln(\Theta) + (nm - \sum x_i) ln(1-\Theta)
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
L(\Theta) =& ln(\prod P(x_i)) \\
\frac{dL(\Theta)}{d\Theta} =& \frac{\sum x_i}{\Theta} - \frac{nm - \sum x_i}{1-\Theta} = 0 \\
\frac{1}{\Theta} =& \frac{nm - \sum x_i}{\sum x_i} + 1 \\
\Theta^* =& \frac{\sum x_i}{nm}
\end{aligned}
\tag{8}
$$

Now to proof that $L(\Theta^*)$ is the maximun probability we must show that $\frac{d^2 L(\Theta)}{d\Theta^2} \mid_{\Theta=\Theta^*} < 0$.

Limit bound Cramer-Rao

a boundarie to limit

Fisher information

Measure the amount of information that have a random variable about a parameter.

concepts about estimators

The quality of estimators:

- **Unbiased estimator :** The mean of the estimator is equal to the mean of parameter.
- **Variance:** the dispersion of the estimations regarding the mean value of the same.
- **quadratic value mean:** offer information about another two.

Topics to research

Teorema rao blackwell Teorema de Basu Teorema de Lehman-Scheffe propiedad de invarianza? teomrea de Zenha

Compare estimators

the lake is shallow

Score Statistics

Suppose that do you have a function $f(x,y) = x^2 + y^2$, then we can define the gradient as follow:

$$\nabla f(x,y) = \begin{bmatrix} \frac{df}{dx} \\ \frac{df}{dy} \end{bmatrix} \tag{9}$$

Notes about gradient

The gradient is perpendicular to each point in the level curve to any surface or level curve.

important resource

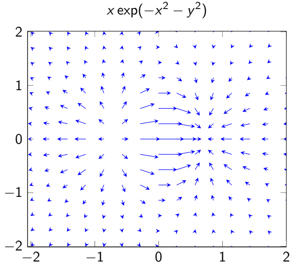
https://www.youtube.com/watch?v=rB83DpBJQsE

Escalar fields

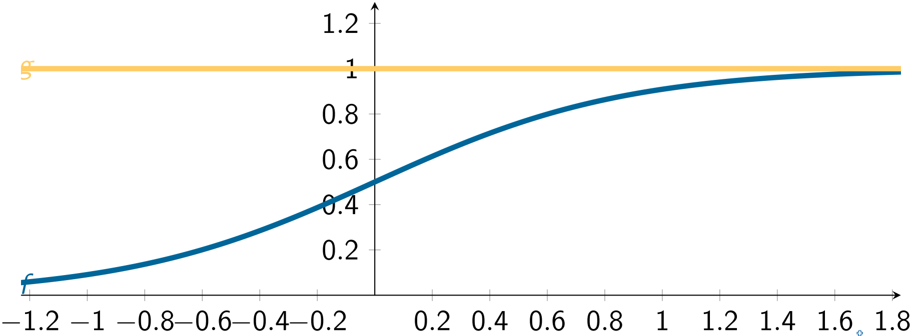
it is a function that $f : R^m \rightarrow R$

Vector field

it is a function $\vec{F} = R^n \to R^n$ that associated a vector to each point in the euclidean space. you can use the gradient to construct the vector field.



Logistic equation



logistic equation

$$f(x) = \frac{\kappa}{1 + e^{-\alpha(x - x_0)}} \tag{10}$$

Where κ it is the maximun value.

logistic equation

Population growth

$$\frac{dy}{dt} = ry(1 - \frac{y}{k}) \tag{11}$$

python implementation

Statsmodels

```
formula = 'died~studytime + C(drug) ' # logit died studytime
    i.drug
model = smf.logit(formula= formula, data=df)
results = model.fit()
print(results.summary())
coefs = pd.DataFrame({
    'coef': results.params.values,
    'odds ratio': np.exp(results.params.values),
    'name': results.params.index
})
coefs
```

Prediction score

The usual way

Talleres

1. structure problems 2. loop problems 3. Combinatorial problems 4. derivates systems 5. matrix algorithm 6. matrix algorithm 7. Economic growth modeling 8. Analytic problem 9. Analytic problem 7 k-means algorithm 8. perceptron algorithm 9. World bank data 10. Icfes data 11. Mortality Data 12. review project

Solve

with newton method solve the following:

$$e^x = 4 - x^2 \tag{12}$$

las dos soluciones son: x1 = 1.05 x2 = -1.96

Entropy

scoring