

# eBox Component API 手册

---

## 打造 eBox 生态圈

版本：版本 19.1.03

发布日期：2019 年 1 月

本指南内容及产如有更新，请参考最新手

勘误记录

日期	摘要	提交者

## 如何使用本 API 手册

本 API 手册只提供了公共基础函数和类中 `public` 函数的解释。并没有对程序中所有的变量和宏作出详细解释。如果涉及到相关宏定义的内容请阅读程序代码中的注释。

本手册更新可能没有程序更新的快，部分内容可能有所书写错误，最终以头文件中的内容为准。

本手册只作为引导性学习使用，请多读程序头文件已获得更全面的信息。

完整的应用示例可参考 `example` 文件夹下的内容

如发现有文字性错误或者部分代码错误请联系 [995207301@qq.com](mailto:995207301@qq.com)。

# 目录

目录.....	1
第 1 章     LinearRegression .....	1
1.1     linear_regression(double *data, int rows, double *a, double *b, double *SquarePoor);.....	1
1.2     multiple_regression (double *data, int rows, int cols, double *Answer, double *SquarePoor); .....	2

# 第1章 LinearRegression

- 描述:

回归分析是对具有因果关系的影响因素（自变量）和预测对象（因变量）所进行的数理统计分析处理。只有当变量与因变量确实存在某种关系时，建立的回归方程才有意义。如果在回归分析中，只包括一个自变量和一个因变量，且二者的关系可用一条直线近似表示，这种回归分析称为一元线性回归分析。如果回归分析中包括两个或两个以上的自变量，且因变量和自变量之间是线性关系，则称为多重线性回归分析。

**回归平方和(SSR):** 因变量的回归值（直线上的 Y 值）与其均值（给定点的 Y 值平均）的差的平方和，即，它是由于自变量 x 的变化引起的 y 的变化，反映了 y 的总偏差中由于 x 与 y 之间的线性关系引起的 y 的变化部分,回归线可以解释的偏差。

**残差平方和(SSE):** 因变量的各实际观测值(给定点的 Y 值)与回归值（回归直线上的 Y 值）的差的平方和，它是除了 x 对 y 的线性影响之外的其他因素对 y 变化的作用，回归直线不能解释的偏差。

**总平方和(SST):**  $SST=SSR+SSE$ ,是每个因变量的实际值（给定点的所有 Y）与因变量平均值（给定点的所有 Y 的平均）的差的平方和，即，反映了 **因变量取值的总体波动情况**。

**判定系数( $R^2$ ):**  $SSR/SST$  或  $1-SSE/SST$ . 表征曲线拟合程度，取值范围 0-1，越接近 1 越好。

假如所有的点都在回归线上，说明 SSE 为 0，则  $R^2=1$ ，意味着 Y 的变化 100%由 X 的变化引起，没有其他因素会影响 Y，回归线能够完全解释 Y 的变化。如果  $R^2$  很低，说明 X 和 Y 之间可能不存在线性关系

关于回归分析的详细解释，请参考 <https://www.cnblogs.com/nxld/p/6123239.html> 或自行百度

## 1.1 linear\_regression(double \*data, int rows, double \*a, double \*b, double \*SquarePoor);

- 描述:

一元线性回归  $Y=A+BX$ ,可以根据样本数据（data）确定常数项 A 和回归系数 B。

- 参数:

double \*data: 指向样本数据 data[rows\*2]数组,数组内容为 X, Y;

int rows: 样本数据行数;

double \*a, double \*b: 返回常数项 A，回归系数 B

double \*SquarePoor: 指向 SquarePoor[4]数组,返回方差分析指标: 回归平方和，剩余平方和，回归平方差，剩余平方差

- 返回值:

0 求解成功, -1 错误

- 例程:

参考 example\math\linearRegression

## 1.2 multiple\_regression (double \*data, int rows, int cols, double \*Answer, double \*SquarePoor);

- 描述:

多元回归方程  $Y = B_0 + B_1X_1 + B_2X_2 + \dots B_nX_n$ , 由多个自变量的最优组合共同来预测或估计因变量

- 参数:

double \*data: 指向样本数据 data[rows\*cols]数组,数组内容为 X, Y;

int rows, int cols: 样本数据行数,列数;

double \*Answer,: Answer[cols], 返回回归系数数组( $B_0, B_1 \dots B_n$ )

double \*SquarePoor: 指向 SquarePoor[4]数组,返回方差分析指标: 回归平方和, 剩余平方和, 回归平方差, 剩余平方差

- 返回值:

0 求解成功, -1 错误

- 例程:

参考 example\math\ multiple\_regression