**MLE**

1. **算法简介**

最大似然估计提供了一种给定观察数据来评估模型参数的方法，即：“模型已定，参数未知”。最大似然估计的一般求解过程如下：

1. 写出似然函数
2. 对似然函数取对数，并整理
3. 对不同参数求偏导
4. 解似然方程

[**http://blog.csdn.net/hezhourongro/article/details/17167717?locationNum=15**](http://blog.csdn.net/hezhourongro/article/details/17167717?locationNum=15)**(该博客讲的更加详细)。**

1. **实例讲解**

假设我们要统计全国人口的身高，首先假设这个身高服从服从正态分布，但是该分布的均值与方差未知。我们没有人力与物力去统计全国每个人的身高，但是可以通过采样，获取部分人的身高，然后通过最大似然估计来获取上述假设中的正态分布的均值与方差。

1. 身高服从正态分布，分布参数为，所以似然函数如下所示：



1. 对似然函数取对数得到如下形式：



1. 对分布参数分别求偏导





上面求偏导前面的常数省略了，因为这不影响第四步的求解。

1. 解释然方程，令第三步得到的偏导为零，可以得到最优的和。

1. **工程文件简介**

src 文件夹下是源代码：

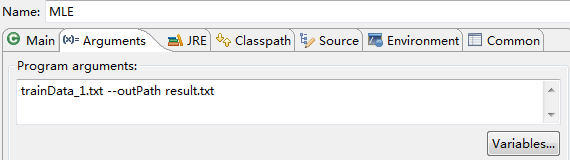
MLEDataProc.java => 读取训练数据。

MLEOption.java => 参数类，该MLE算法支持的参数。

MLE.java => 求解和。

1. **程序调用方式**

* 在**windows**下可以用eclipse打开该工程，在Run =>Run Configurations界面设置程序需要的参数，参数细节在后面会给出解释。



* 在**Linux** 下可以可以通过如下命令进行调用。

把文件夹下的MLE.jar 拷贝到linux下某个目录，通过如下命令进行调用。

java –jar MLE.jar trainDataFilePath [--outPath outPathValue]

**方括号中表示可选参数。**

trainDataFilePath：表示训练数据文件路径

outPath：把分布参数最优值输出到文件中。如果不传，默认为空，则输出到终端。

1. **数据集**

在文件夹下有一个数据集文件trainData\_1.txt。该文件记录了一部分人的身高。

程序的时间复杂度是,其中n表示数据集的大小。