教育技术应用

基于 Python 和梯度下降算法的物理实验数据一元线性拟合方法

关毅铭　程敏熙

（华南师范大学物理与电信工程学院　广东 广州　510006）

（收稿日期：2019－03－04）

摘　要：介绍了用Python实现梯度下降算法的方法. 以3个大学物理实验（光电效应测量普朗克常量、光泵磁共振测量朗德因子以及偏振光实验验证马吕斯定律）为例，用相关系数 r 和拟合优度 R^2 作为指标，检验梯度下降算法对物理实验数据线性拟合处理的效果，对以 Python 和梯度下降算法为切入点，了解深度学习和人工智能算法有启发意义.

关键词：Python　梯度下降法　线性拟合　大学物理实验

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. 梯度下降法是一种最优化算法，常用于机器学习和人工智能中用来递归性地逼近最小偏差模型，对数据量大，关系复杂，很难得到解析解的函数关系来说十分有效[1]. 它是迭代法的一种，可以用于求解最小二乘法问题. 所以我们也可以用梯度下降法来处理呈线性关系的实验数据.

1　梯度下降法的介绍

1.1　梯度下降法

一元线性回归其实就是去找到一条直线，这条直线能以最小的误差来拟合数据.

如图1所示，横坐标表示 x，纵坐标表示 y. 一元线性回归要找的就是图中的这条直线，用 $y=b+mx$ 表示. 但是，所有的点都在这条拟合直线上肯定不现实，所以这些点要尽量靠近这条直线. 即去找每个点和直线的距离 $|y_{ie}=b+mx_i|$ 最小的那条线. 为了简单起见，将绝对值转化为平方，那么误差可以表示为

$$\text{Loss}=\frac{1}{N}\sum_{i=1}^{n}(y_{ie}-y_i)^2=\frac{1}{N}\sum_{i=1}^{n}[y_i-(b+mx_i)]^2$$

这里 i 表示第 i 个数据，N 表示总的样本个数. 把这个损失方程看作是 m 和 b 的方程. 作为一个 m 和 b 的二次方程，那么求损失函数最小值的问题就转变成了求极值问题. 我们可以用梯度下降法求解 m 和 b 的值以确定拟合直线.

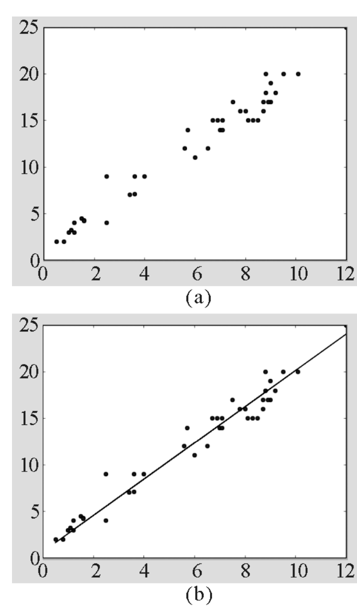


图1　一元线性回归拟合原理

为了更好地理解梯度下降法，这里把损失函数

作者简介：关毅铭（1996－　），男，在读硕士研究生，研究方向为学科教学（物理）.

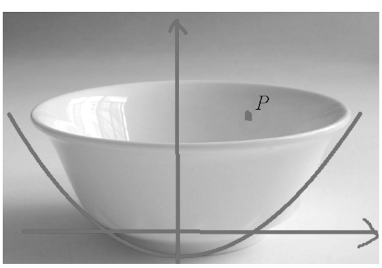
通讯作者：程敏熙（1962－　），男，博士，副教授，研究方向为光电技术与系统、物理实验设计.

想象成一个碗，而函数的最小值就是碗底，如图2(a)所示．不管位于这个碗的什么位置，只要往下走，也就是往梯度方向走，沿着梯度一点一点滑下去，就能达到这个最低点．梯度就是下面两个式子．

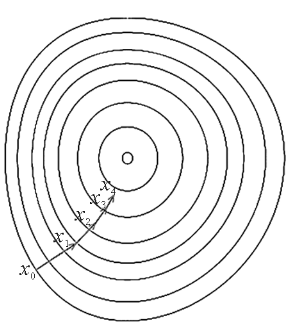
$$\frac{\partial \mathrm{Loss}}{\partial m}=\frac{\partial\{\frac{1}{N}\sum_{i}^{n}[y_i-(m+bx_i)]^2\}}{\partial m}=-\frac{2}{N}\sum_{i=1}^{n}x_i[y_i-(b+mx_i)]^2$$

$$\frac{\partial \mathrm{Loss}}{\partial b}=\frac{\partial\{\frac{1}{N}\sum_{i}^{n}[y_i-(b+mx_i)]^2\}}{\partial b}=-\frac{2}{N}\sum_{i=1}^{n}[y_i-(b+mx_i)]^2$$

接下来就要定义步长，用来表示每次滑多长，再定义一个迭代值用来表示滑多少次，这样就能一点一点地靠近最小值．定义好这两个值，计算机就可以一边求梯度，一边向下滑，去更新 m 和 b，最后得到损失函数最小时对应的 m 和 b 的值．



(a)



(b)

图 2　梯度下降法示意图

也可以这样理解梯度下降法．如图2(b)所示，每一个圈代表一个函数梯度，最中心表示函数极值点，每次迭代根据当前位置求得的梯度(用于确定搜索方向以及与步长共同决定前进速度)和步长找到一个新的位置，这样不断迭代最终到达目标函数局部最优点(如果目标函数是凸函数，则到达全局最优点)．

实现梯度下降法的主要Python代码如下所示．

```
def optimizer(starting_b,starting_m,learning_rate,num_iter):
    b = starting_b
    m = starting_m
    for i in range(num_iter):
        b,m = compute_gradient(b,m,learning_rate)
    return [b,m]
def compute_gradient(b_current,m_current,learning_rate):
    b_gradient = 0
    m_gradient = 0
    N = float(len(data))
    b_gradient = -(2/N) * (y-m_current * x - b_current)
    b_gradient = np.sum(b_gradient,axis= 0)
    m_gradient = -(2/N) * x * (y-m_current * x - b_current)
    m_gradient = np.sum(m_gradient,axis=0)
    new_b = b_current - (learning_rate * b_gradient)
    new_m = m_current - (learning_rate * m_gradient)
    return [new_b,new_m]
```

1.2　检验拟合结果的指标

求出 m 和 b 的具体数值，即找到了拟合的直线．一般数据处理中会用相关系数 r 检验两变量的相关性，用拟合优度 R^2 检验回归直线的拟合程度．

相关关系是一种非确定性的关系，相关系数是研究变量之间线性相关程度的量．相关系数又叫线性相关系数，一般用字母 r 表示，用来度量两个变量间的线性关系．当 r 接近于 ± 1 时，可以说两个变量相关性比较好．

相关系数公式为[2]

$$r=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2}}$$

拟合优度是指回归直线对观测值的拟合程度．度量拟合优度的统计量是可决系数 R^2．R^2 最大值为

1. R^2 的值越接近 1，说明回归直线对观测值的拟合程度越好；反之，R^2 的值越接近零，说明回归直线对观测值的拟合程度越差.

拟合优度公式为[3]：

$$R^2=\frac{SSR}{SST}=1-\frac{SSE}{SST}=1-\frac{\sum(y_i-\hat{y})^2}{\sum(y_i-\bar{y})^2}$$

2　用 Python 处理 3 个物理实验数据

2.1　光电效应测普朗克常量[5]（5 个点）

根据爱因斯坦的光电效应方程

$$\frac{1}{2}mv^2=eU_S$$

可以推导出普朗克常量

$$h=\frac{e(U_2-U_1)}{\nu_1-\nu_2}$$

根据实验数据拟合出遏止电压和频率的二元一次方程，求出斜率 k，便可求得普朗克常量 $h=ek$.

实验数据和运用 Python 实现梯度下降法的结果，如表 1 和图 3 所示.

表 1　光电效应测普朗克常量实验数据

波长 λ/nm	365.0	404.7	435.8	546.1	577.0
频率 /($\times 10^{14}$ Hz)	8.219	7.413	6.884	5.493	5.199
遏止电压 U_0/V	−1.87	−1.41	−1.18	−0.66	−0.56
斜率 $k\approx 0.417\,9$	普朗克常数 $h\approx 6.686\times 10^{-34}$ J·s		相对误差 $\delta=0.92\%$		

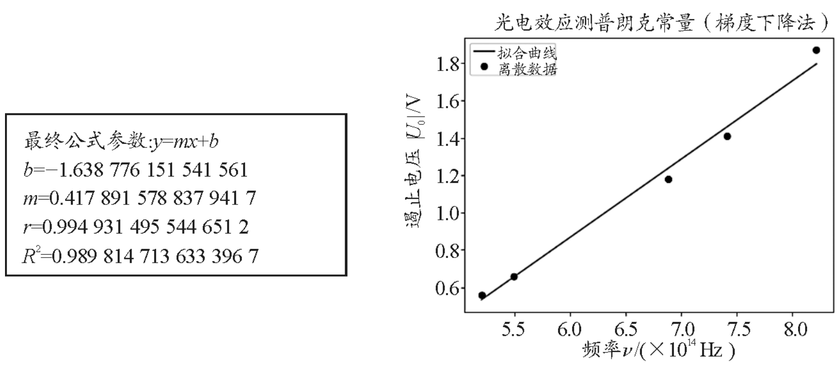


图 3　Python 实现遏止电压和频率一元线性拟合

2.2　光泵磁共振测朗德因子 g_F[4]（10 个点）

光泵磁共振实验中，共振条件为

$$h\nu=\Delta E=g_F\mu_B B_0$$

此式可改写为

$$\nu=\frac{g_F\mu_B B_0}{h}$$

显然，ν 和 B 之间为线性关系. 即使考虑地磁场水平分量和扫场直流分量的影响，拟合直线的斜率也不会改变. 设拟合的直线方程为

$$\nu=a_0+a_1B$$

则斜率

$$a_1=\frac{\mu_B g_F}{h}$$

从而

$$g_F=\frac{ha_1}{\mu_B}$$

可见只要求出直线斜率 a_1，代入上式便可求得 g_F 的值.

实验数据和运用 Python 实现梯度下降法的结果，如表 2 和图 4 所示.

表 2　光泵磁共振测 ^{85}Rb 朗德因子实验数据

水平磁场 /($\times 10^{-5}$ T)	7.01	7.94	8.88	9.81	10.7
频率 ν/kHz	603.7	647.3	690.0	731.3	780.3
水平磁场 /($\times 10^{-5}$ T)	11.7	12.6	13.6	14.5	15.4
频率 ν/kHz	830.1	872.3	917.6	959.8	1 007.3
斜率 $a_1\approx 48.13$	^{85}Rb 朗德因子 $g_F\approx 0.344$		相对误差 $\delta=3.2\%$		

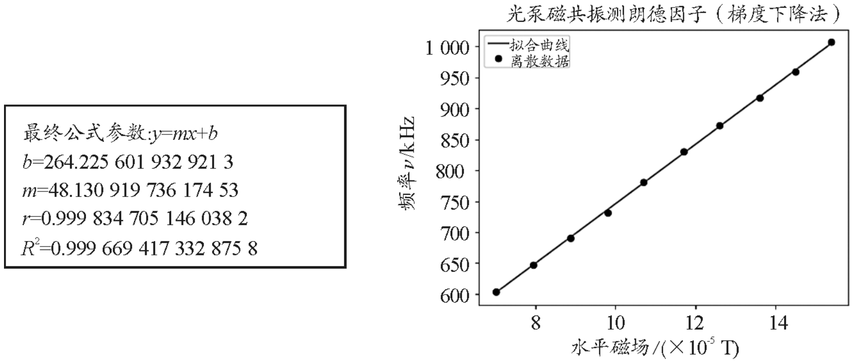


图 4 Python 实现频率和水平磁场一元线性拟合

2.3 偏振光实验验证马吕斯定律[5](16 个点)

马吕斯定律为$I_\theta = I_0 \cos^2\theta$. 显然,$I_\theta$ 和 $\cos^2\theta$ 为线性关系. 大学物理实验中一般用偏振光实验验证马吕斯定律.

实验数据和运用 Python 实现梯度下降法的结果,如表 3 和图 5 所示.

表 3 偏振光实验验证马吕斯定律实验数据

$\cos^2\theta$	0	0.010 93	0.043 23	0.095 49	0.165 43	0.25	0.345 49	0.447 74
$I_\theta/(\times 10^{-7}$ A)	0	9.5	39	94	170	253.5	368	486.5
$\cos^2\theta$	0.552 26	0.654 51	0.75	0.834 57	0.904 51	0.956 77	0.989 07	1
$I_\theta/(\times 10^{-7}$ A)	611.5	733.5	852.5	960	1 043	1 106	1 143.5	1 174
斜率 $k = I_0 \approx 1\ 172 \times 10^{-7}$ A					相对误差 $\delta = 0.17\%$			

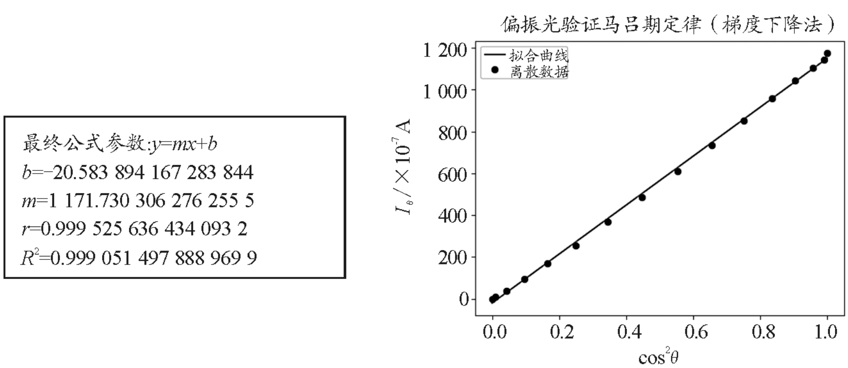


图 5 Python 实现 I_θ 和 $\cos^2\theta$ 一元线性拟合

3 结果分析和结论

本研究在实现梯度下降法时,设置步长为 0.001,迭代值为 75 000. 我们可以根据实际情况,调大步长或调小迭代值,从而提高计算速度.

光电效应求普朗克常量中,Excel 得出的拟合优度 $R^2 = 0.989\ 9$,梯度下降法得到的拟合优度$R^2 = 0.989\ 8$;光泵磁共振测朗德因子中,Excel 得出的拟合优度 $R^2 = 0.999\ 7$,梯度下降法的拟合优度 $R^2 = 0.999\ 6$;验证马吕斯定律中,Excel 得出的拟合优度 $R^2 = 0.999\ 1$,梯度下降法得到的拟合优度 $R^2 = 0.999\ 0$. 可见,Excel 的 R^2 和梯度下降法的 R^2 结果基本一致. 相关系数 r 和拟合优度 R^2 在保留有效数字时,通常保留 4 位,而且只舍不入[7]. 而 Excel 的结果默认四舍五入,所以两者结果其实是一样的.

梯度下降法与 Excel 的最小二乘法相比,梯度下降法需要选择步长,而最小二乘法不需要. 梯度下降法是迭代求解,最小二乘法是计算解析解. 物理实验数据处理中,样本量不算很大,梯度下降法与最小二乘法的计算速度大致相同,结果也相同. 采用 Python 和梯度下降算法对我们了解深度学习和人工智能研究及处理相关数据有启发意义.

信息技术在磁场对通电导线作用力研究中的应用

严　涵
（南通大学附属中学　江苏 南通　226001）

（收稿日期：2019-03-13）

摘　要：基于移动设备的增强现实技术（AR）的发展，可以直接在实验装置图片上进行虚拟物理量的叠加，从而加深学生对实验现象的内涵理解.

关键词：信息技术　物理课堂　融合式教学

当前基于“培养核心素养，提升关键能力”的新一轮教学改革正如火如荼的进行着.高中物理所教知识体系基本没有变化，但是如何在不变中，结合当代学生学习特点，实践出一条符合时代要求的高效物理课堂教学法，需要教师不断地摸索与尝试.在信息技术日新月异发展的过程中，如何结合新媒体技术与传统物理课堂教学，让两者有机融合，笔者做了一定的尝试.下面就以人教版选修 3-1“磁场”一章中的“磁场对通电导线的作用力”课堂实录为例，谈谈信息技术与物理课堂教学的融合思路.

本节知识需要学生知道什么是安培力.知道左手定则是判断通电导线在磁场中所受安培力方向的基本方法.但由于左手定则的三维空间感要求高，学生不容易掌握，因此采用虚拟现实技术让抽象变成

参 考 文 献

1 斋藤康毅.深度学习入门：基于 Python 的理论与实现.北京：人民邮电出版社，2018.102 ～ 119
2 R. Taylor. An Introduction to Error Analysis (SECOND EDITION). California: University Science Books，1996.221
3 李子奈，叶阿忠.高等计量经济学.北京：清华大学出版社，2000.260 ～ 263
4 嵩天，礼欣，黄天羽.Python语言程序设计基础（第 2 版）.北京：高等教育出版社，2017
5 吴先球，熊予莹.近代物理实验教程（第 2 版）.北京：科学出版社，2009.295 ～ 303
6 马颖，梁鸿东，徐丽琴.大学物理实验教程（第 2 版）.北京：清华大学出版社，2013.73 ～ 79，191 ～ 195
7 中国国家标准化管理委员会.中华人民共和国国家标准：GB/T5750.3—2006.北京：中国标准出版社，2006

Linear Fitting Method of Unitary about Physical Experimental Data Based on Python and Gradient Descent Algorithm

Guan Yige　Cheng Minxi
(School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou, Guangdong　510006)

Abstract: A method of implementing gradient descent algorithm in Python was proposed. With three college physics experiment (Photoelectric effect measured Planck constant, optical pump magnetic resonance measured Lander factor and polarized light verified Malus's Law) as examples, using the correlation coefficient r and goodness of fit R^2 as indicators, the effect of the linear fitting process on the physical experimental data by the gradient descent algorithm was verified. It is instructive to understand deep learning and artificial intelligence algorithms with the perspective of Python and gradient descent algorithm.

Key words: Python; gradient descent; linear fitting; college physics experiment