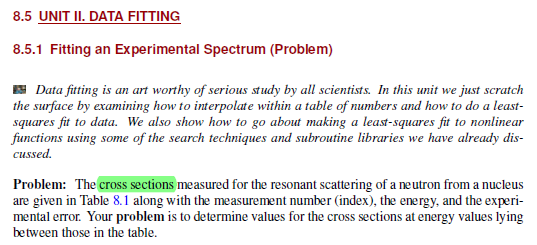
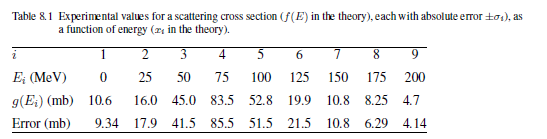
**2021**春季研究生计算物理**II**期末作业

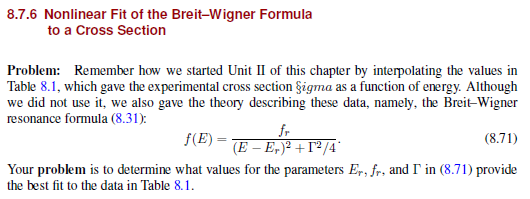
2020102020001 何显龙 2021.06.29

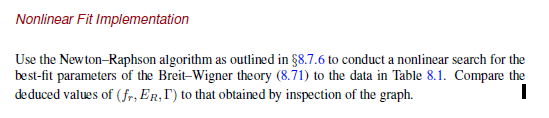
## 题目：

P180节8.5 **UNIT II. DATA FITTING**中**8.5.1 Fitting an Experimental Spectrum (Problem)**的问题，具体回答P193中8.7.6节末的问题Nonlinear Fit Implementation









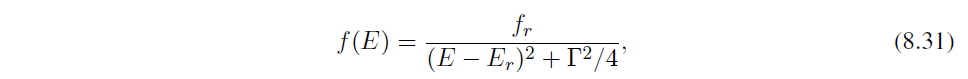
## 问题：

原子核对中子的共振散射实验中，在不同入射中子能量（速度）Ei下，测量到的散射截面g(Ei)数据如表8.1所示，一并给出每次测量误差。我们的问题是，根据已知的测量数据，推断出在实验得到的入射能量范围之间、其他未做实验的能量处，散射截面的具体值。

这是一个非常典型的自然科学方法实验，即在受各种实际因素限制的情况下，使用有限的、已知的探测数据点，基于某些假设和理论，去估计、预测无限多（无限小间隔）、未知的未来可能的实验结果。

有很多方法可以解决这一问题：有一大类是纯现象学的插值方法，即在各实验数据点的间隔中以特定方式插入估计值，虽然直接且相对简单、但完全忽视了噪声（误差）。另一大类是基于特定理论的参数拟合，即通过某些理论（如第一性原理出发）直接计算自变量、因变量之间的函数形式（即依赖关系），其中涉及的参数由实验数据反推计算给出（限制），这种。当然，很多时候在理论基础相当不完备的时候，常常也会先观察数据走势、甚至在插值函数的基础上，提出所谓带参数的“现象学”理论模板，再用后者进行参数拟合（估计）。

按照本文倾向于相信“正确”的Breit–Wigner理论公式：



及P194的说法，式中fr表征峰值高度（实际高度为4\*fr/Γ^2），Er为峰值能量，Γ为半高全宽（full width at half-maximum）。最终目的是参照P193节8.7.6，基于表8.1中的数据，使用Newton–Raphson算法（也称牛顿迭代法），使用Breit–Wigner公式做非线性最优参数估计，并比较这种计算结果与直接观察散点图的估计（这一估计可用作迭代初值的赋予）。

## 解答：代码使用Jupyter Notebook，基于Python环境

### 估计值/迭代初值：

作为相对简单的现象学插值法，主要有两种方式：**Lagrange**插值、三次样条插值。由于其不是本题目的重点，在此仅简略介绍回顾。Lagrange插值的核心思想，是使用Taylor展开多项式对全局global插值，即假设未知函数的依赖公式可以展开为有限阶多项式，其中展开系数待定，每经过的一个实验数据点可以建立一个方程，联立反解展开系数即可。但这一方法在数据点较多时，其拟合结果很容易出现非物理的振荡（Runge现象）、仅在局域local结果可信。

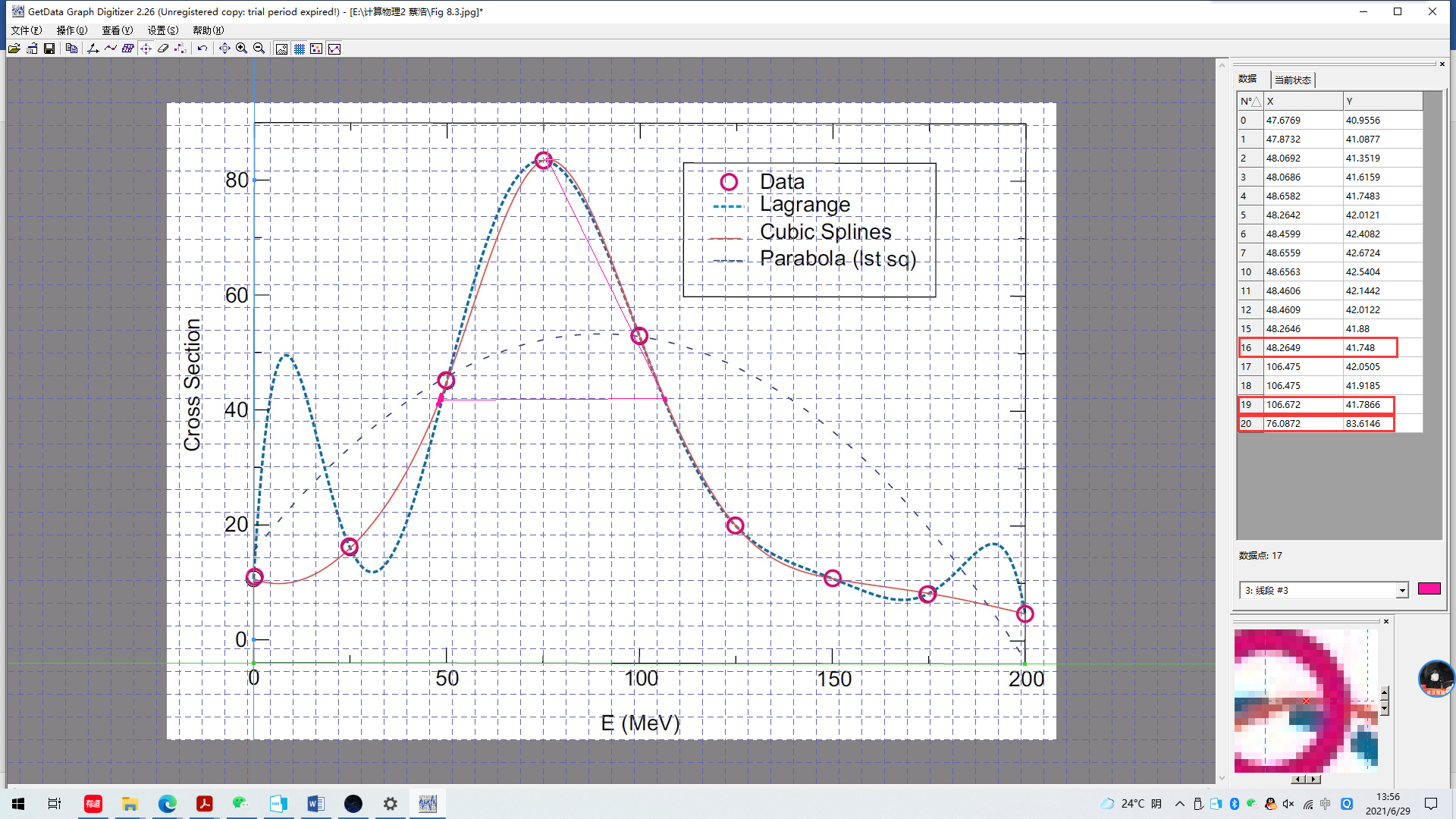
为避免，实际广泛应用的是局域local的三次样条插值。对最简单的一次样条插值，相当于两点之间使用直线段刚性连接，等效于一次Lagrange插值。很容易发现，这种总分段函数、数据点左右导数不相等；为平滑，至少需要再提高一次幂，以使一阶导数连续；通常的还需求二阶导数连续，因此使用（分段）三次函数作为样条工具去插值。

基于这种相对简单的现象学插值，相较原始的仅有散点数据，可以读出更丰富的信息。突出表现在可以直接计算出峰值位置、高度、半高全宽的具体值，在下一步更准确的Breit–Wigner理论公式的参数估计中，以此计算值、而非散点图估计值，作为迭代初值输入，能更合理、相对更快达到目标精度。

由于本题目重点不在于插值，此处仅使用三次样条插值结果的图，以按题目要求“that (the

best-fit parameters) obtained by inspection of the graph”或“guess from a graph of the data.”，从图片（散点图）中“检测inspection”出的结果初值输入并比较。

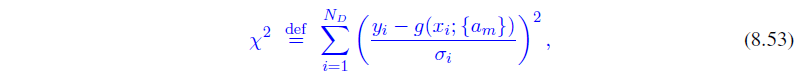
使用getdata软件，直接读取P182图8.3中Cubic Splines曲线的结果，观察得峰值位置X=76.0872，Y=83.6146，半峰高41.8073，半峰全宽ΔX=106.672-48.2649=58.4071。



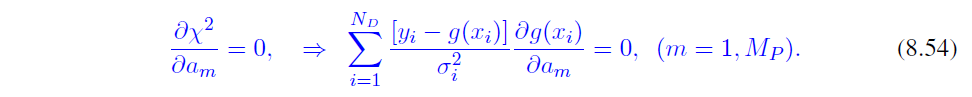
### χ方及最小二乘法：

前文提及的插值法，最关键的是插值函数需要经过所有的点。而另一种“拟合”方法，则是寻求全局近似，不求经过每一点。最小二乘法是对拟合效果的最常用判据，即要求理论值偏离实验值的方差最小。反解出来的参数，就可以认为是最佳拟合的参数曲线。这在概率学上被称作“参数的点估计”。

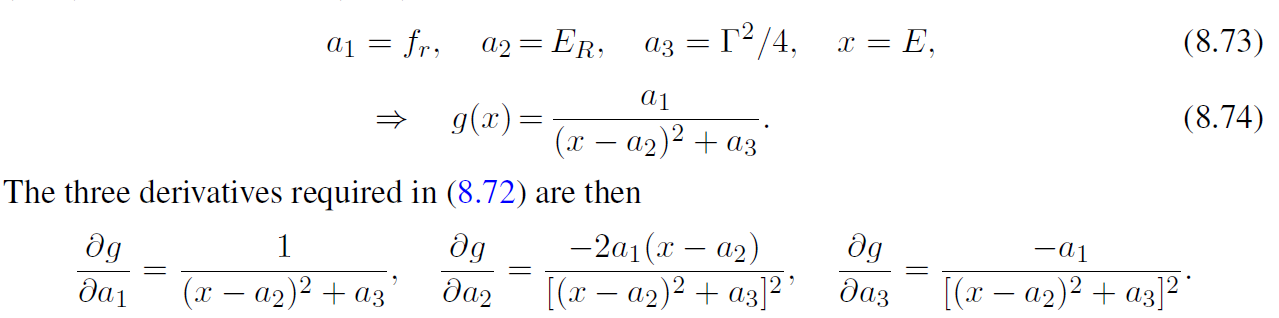
最早、最简单的案例是直线拟合、或称为一元线性回归，对非线性函数的最小二乘法可以说是基于此发展而来。通常我们定义chi方作为判别函数（待估计参数的函数），其越小表示使用这一参数估计的越好：其中负号代表偏离量，平方代表正负偏离平权；权重因子σ表示，方差大、测量不够准确的数据对整体误差函数影响小；总体求和表示考虑多次测量。其更自然的数学解释，源自概率论中的极大似然估计。



线性回归在计算上之所以简单，是因为函数依赖形式仅一次幂，χ方仅（两个未知参数的）二次函数，十分容易分析其极值位置、甚至χ方的全局趋势，即可以直接令其偏导等于0，解析计算其最佳拟合值。甚至只要是多项式拟合、高次幂都存在原则上的解析解。



而对于如题的非线性函数的拟合，原则上令其偏导为0、也可以计算极值，只不过得到的是非线性矩阵方程，数值上常用牛顿迭代法解。



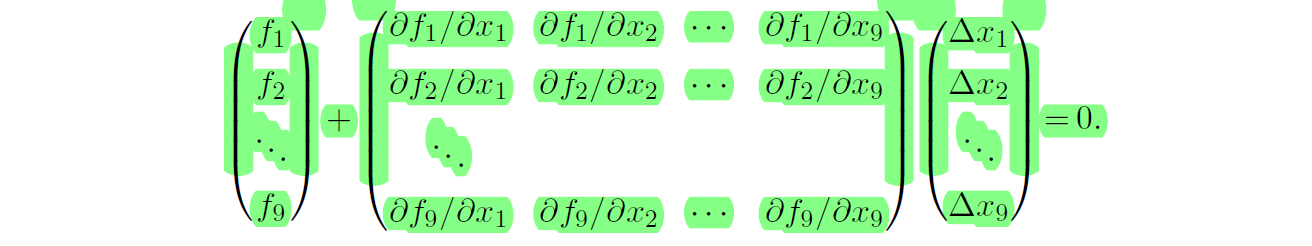
书P194式8.75-8.79中，不清楚为何方差因子不存在了，在此重新加上，否则相当于认为，每次测量值误差都完全相等。同时注意到表8.1中，误差与测量值竟然十分接近！参考反应截面的典型测量值：

<https://download.csdn.net/download/weixin_38717143/19404087?spm=1001.2101.3001.5697>，将表格中所给误差手动乘以0.1，以用测试。



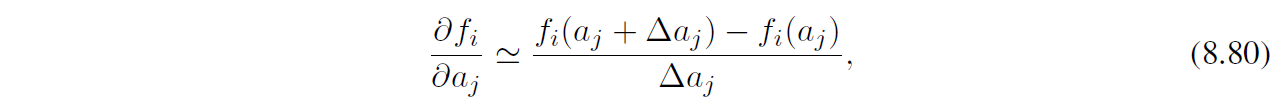
### 牛顿迭代法：

为求解非线性方程组，将应用于单个非线性方程的牛顿切线迭代法加以推广，核心公式如下：



即为求解f(ai)=0，先取试探解ai，再令下一步真值f(a+Δa)=0在ai点的泰勒展开，转化为计算线性方程组的解Δa，最后反复迭代至误差满足精度要求。参照前文三次样条插值的参数估计值，a0=(70786.3,76.0872,852.847)。

对于本题，由于fi(ai)已经是待拟合非线性函数g(x,ai)对ai的一阶导数、略显复杂，再对其求一次偏导会显得更加繁琐，不利于输入与查错。因此参照教材建议，使用嵌套联结的向前微分近似：



### 线性方程组：

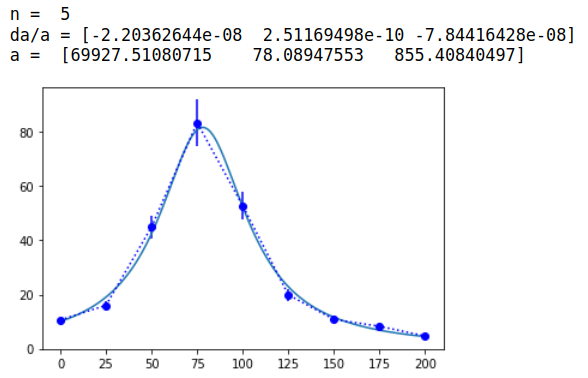
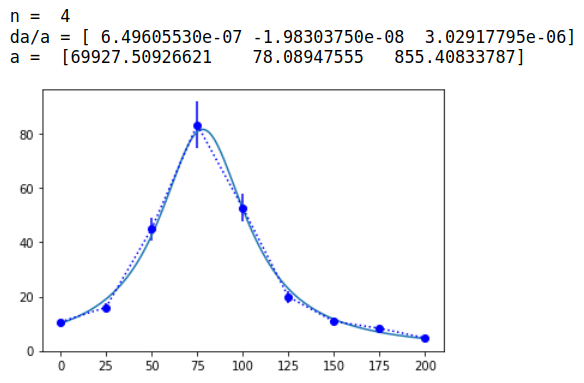
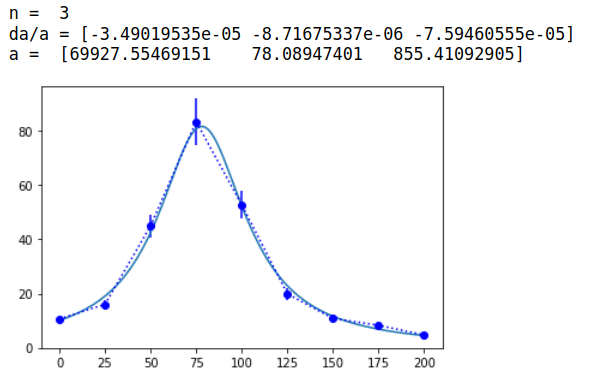
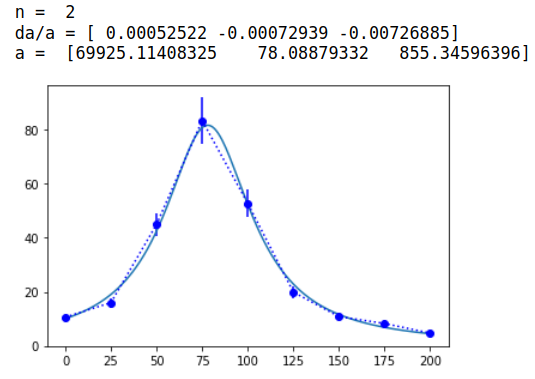
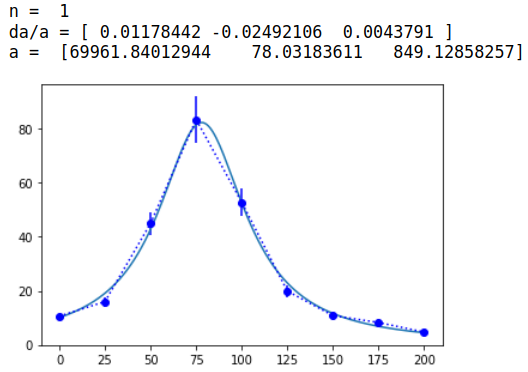
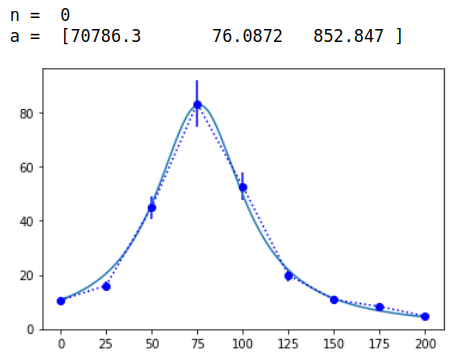
最终，我们将参数估计的问题，转化为迭代求解线性方程组的问题。此处不是本题目的研究重点，因而直接引用P173提及的Python数学包Numpy中的线性代数包linalg。

### 结果

退出迭代循环的判别条件，本题设定为，三个参数相对误差的标准差≤1.e-6。

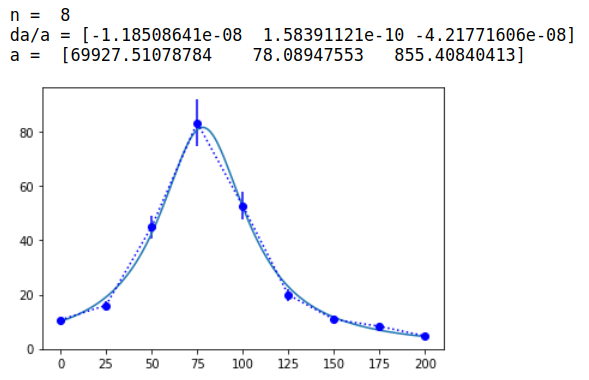
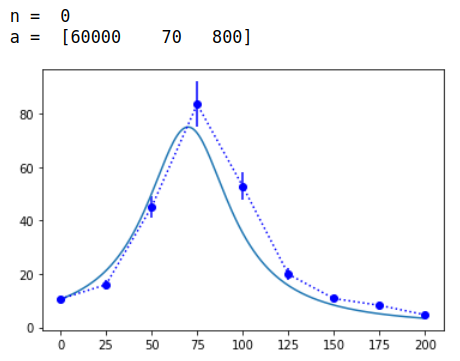
最终结果图以带误差棒的实验数据散点图，与参数估计值、及估计分布图形式展示。

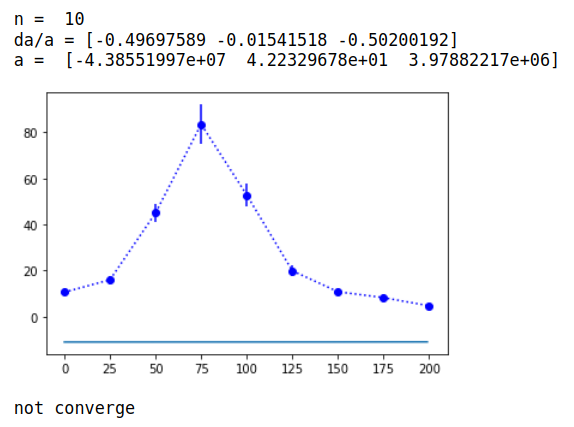
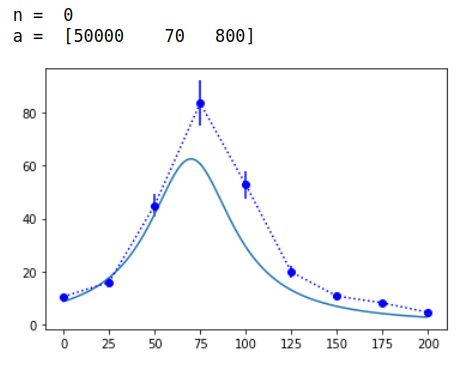
以下从左至右、从上至下，分别为迭代0次（三次样条插值结果，作为初始迭代输入值）、1次、2次、……，仅5次迭代就能达到极高的精度！



### 讨论：

容易看出，牛顿切线迭代法，从目标值附近的初值迭代，收敛速度相当快；但如下其他测试值发现，若偏离目标值较远，很容易发散。





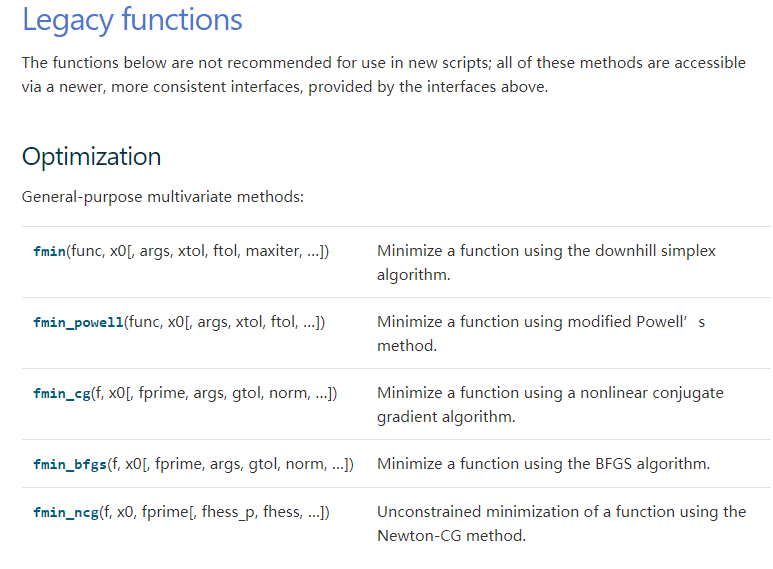
推测是因为牛顿切线法很难保证，多个参数在非线性依赖下，同时取到其极值位置，很容易顾此失彼导致迭代发散。仅在零点附近，才具有较快、较稳定的收敛速度，尤其适用于低维。

### 下一步

为求解判别函数chi方极值点给出的非线性方程，还可以考虑“一般（不动点）迭代法”尝试。进一步，求解这一非线性方程的目的，实质上是为了找到chi方的极小值点，“最速下降法”是针对这一问题相当常用的一个算法。考虑到作业篇幅，此处不再尝试，目前已有成熟的Python包：from scipy.optimize import fmin

模块<https://docs.scipy.org/doc/scipy/reference/optimize.html#module-scipy.optimize>

利用Nelder-Mead 法或称下山单纯形法，计算函数最小值。



更进一步，如果我们还希望知道chi方除极值点之外，所有参数空间可能的信息，原则上还可以遍历全（局域）参数空间，计算chi方的分布函数，即所谓corner图，再从中观察最合适的参数估计值。

事实上，这一思路正是所谓“贝叶斯统计学”的标准模式。即把未知参数也视作随机变量，服从某种分布。这样chi方就有具体的物理意义了，描述的是，假设多次（每次）实验测量服从高斯分布（大数定理的基础下），那么chi方就是在给定参数（条件概率）后、得到测量值的概率的联合密度分布（的负对数）。最小二乘法的更深层理解其实是，均匀先验下的贝叶斯后验众数估计。基于此参数的后验分布，还可以给出实验结果的后验预测分布。

此例的非线性依赖关系尚且稍简单，遇到更复杂依赖、参数更多的情况，计算最小值、遍历参数空间的点都稍显臃肿。更常用的是MCMC方法，即在未知全局chi方函数分布图像的基础上，仍能按其分布做蒙特卡洛抽样，并遵循马尔科夫链一步步抽取下去，直到分布轮廓满足精度需求。这正是我目前科研“参数估计宇宙学”这一大领域的主要技术手段，限于篇幅，在此不做展示。