问题一要建立双峰复合光谱的分峰模型，并给出各峰的中心波长、高度及半峰宽。先对附件1中的三组数据分别绘制双峰复合光谱信号强度与波长的散点图，结合数据与图像便可粗略估计各自对应的单峰光谱波长、峰高及半峰宽数值，从而大致确定各单峰曲线的位置和参数。再分析荧光光谱的形成过程，可得理想状态下单峰光谱的峰强与波长满足一维高斯分布，故采用高斯函数对单峰光谱进行拟合。对于每一个双峰复合光谱，将其对应的两个单峰的预估波长、峰高及半峰宽数值代入高斯函数，便可得到两个较为粗略的单峰分峰。再利用最小二乘法原理，使两个预估分峰的叠加图与实际复合图之间的误差平方和达到最小，便可较为准确地实现双峰复合光谱的分峰，进而求出分峰后的独立光谱曲线与原始光谱曲线的线性相关系数。

问题二要在问题一的基础上探究峰高比对双峰分峰模型的影响。可以在满足瑞利判据的前提下构建虚拟数据，对给定的波长和半峰宽，令峰高比从1.000以0.333步长遍历至9.000，大致模拟实际光谱峰高比范围。将构建的虚拟数据输入问题一的双峰分峰模型，比较拟合图与数据图的线性相关系数与相对误差，从而分析峰高比对双峰分峰效果的影响。

11

利用附件 2 中三峰和四峰叠加的光谱曲线及数据，建立多个光谱峰叠加的复合光谱曲  
线模型，将多峰复合光谱曲线拆分为独立光谱曲线（最多四个峰），并给出分离后各峰的中心  
波长、高度及半峰宽，与其相应的准单色光谱进行比

问题三要建立三峰和四峰复合光谱的分峰模型，并给出各峰的中心波长、高度及半峰宽。此问相较于问题一仅发生了峰数变化，可尝试套用双峰分峰思路建模。注意到多峰复合光谱的重叠度较双峰更高，分析数据图像得到的波长、峰高及半峰宽估计值较真实值的偏差会更大，但初始估计值的偏差可以通过最小二乘法的迭代步数抵消，误差的平方和仍会达到最小。故仍然可对附件2中的八组数据分别绘制复合光谱信号强度与波长的散点图，粗略估计各自对应的单峰光谱波长、峰高及半峰宽。再用高斯函数拟合各单峰曲线，使用最小二乘法迭代至拟合叠加图与实际复合图的误差平方和最小，便可较为准确地实现多峰复合光谱的分峰，进而求出分峰后的独立光谱曲线与原始光谱曲线的线性相关系数。

为了得到峰高的初步估计值，可使用python中的signal.find\_peaks函数在信号强度数据中寻找极值点。考虑到数据受噪声影响导致极值点较多，但噪声对光谱峰信号的干扰远小于峰体本身的信号强度，不会改变光谱峰的大致形状，故可以对极值点的信号强度范围做限制，筛选出双峰复合光谱波峰与波谷处对应的极值点。将复合光谱两个波峰处的信号强度作为分峰的峰高估计值。再使用数据索引得到波峰处对应的波长，作为两个分峰得中心波长估计值。

观测散点图可知，波长小于400nm或大于700nm范围内三组数据的信号强度均较为平坦，相较于波峰的信号强度变化不大。故对每组前100个数据的信号强度取平均数，用来表示峰底的信号强度。再将两个波峰的信号强度分别与峰底的信号强度取平均值，便可得到两个分峰的半峰高。继而检索信号强度高于半峰高的数据点，提取其对应波长的最小值与最大值

大值

数值迭代法分为两类：一类是导数法，该方法以目标函数导数的负方向作为下一次迭代方向；另一类是直接法，该方法一般是在目标函数的导数难以得到的情况下，直接利用目标函数而不利用其导数来确定下一次迭代方向。一般而言，前者的计算效率比后者高。本文目标函数Ｅ的导数存在并较易得到，故采用导数法。

PCR （聚合酶链式反应）是一种用于放大扩增特定DNA片段的分子生物学技术，近年来广泛应用于核酸检测，可以大大提高检测的准确性和应用性。荧光定量PCR技术在PCR基础上引入了荧光基团，可以实时检测产物中的荧光信号并进行定量分析。实际在荧光信号的检测过程中，不同荧光基团产生的荧光具有不同的光谱范围，需要对荧光光谱进行区分。实现复合光谱分峰的方法，从采集的复合光谱曲线中分离不同荧光基团对应的光谱曲线，能够在保证灵敏度的同时对光谱进行区分，对于核酸检测技术及荧光定量PCR反应技术都具有重要意义。

针对问题一：先对附件1中的三组数据分别绘制双峰复合光谱信号强度与波长的散点图，结合数据与图像便可粗略估计各自对应的单峰光谱波长、峰高及半峰宽数值，从而大致确定各单峰曲线的位置和参数。再分析荧光光谱的形成过程，可得理想状态下单峰光谱的峰强与波长满足一维高斯分布，故采用高斯函数对单峰光谱进行拟合。对于每一个双峰复合光谱，将其对应的两个单峰的预估波长、峰高及半峰宽数值代入高斯函数，便可得到两个较为粗略的单峰分峰。再利用最小二乘法原理，使两个预估分峰的叠加图与实际复合图之间的误差平方和达到最小，便可较为准确地实现双峰复合光谱的分峰，进而求出分峰后的独立光谱曲线与原始光谱曲线的线性相关系数。

针对问题二：问题二要在问题一的基础上探究峰高比对双峰分峰模型的影响。可以在满足瑞利判据的前提下构建虚拟数据，对给定的波长和半峰宽，令峰高比从1.000以0.333步长遍历至9.000，大致模拟实际光谱峰高比范围。将构建的虚拟数据输入问题一的双峰分峰模型，比较拟合图与数据图的线性相关系数与相对误差，从而分析峰高比对双峰分峰效果的影响。

针对问题三：此问相较于问题一仅发生了峰数变化，可尝试套用双峰分峰思路建模。注意到多峰复合光谱的重叠度较双峰更高，分析数据图像得到的波长、峰高及半峰宽估计值较真实值的偏差会更大，但初始估计值的偏差可以通过最小二乘法的迭代步数抵消，误差的平方和仍会达到最小。故仍然可对附件2中的八组数据分别绘制复合光谱信号强度与波长的散点图，粗略估计各自对应的单峰光谱波长、峰高及半峰宽。再用高斯函数拟合各单峰曲线，使用最小二乘法迭代至拟合叠加图与实际复合图的误差平方和最小，便可较为准确地实现多峰复合光谱的分峰，进而求出分峰后的独立光谱曲线与原始光谱曲线的线性相关系数。

针对问题四：要建立非对称复合光谱的分峰模型。利用题目数据绘图可知，非对称复合光谱各峰重叠度极大，曲线形状较不规则，无法用前三问的方法直接从原数据中预测各分峰的位置及参数。鉴于理想状态下单峰光谱的峰强与波长满足一维高斯分布，不妨假设非对称光谱可由两个半峰宽不同的高斯曲线各取一半而拼成，进而对非对称曲线实现参数化描述。采用小波分析与二次微分相结合的方法处理数据，降低数据噪声并提取分峰参数的的估计值，再用最小二乘拟合的方法逼近期望输出，实现非对称复合光谱的分峰。

\begin{enumerate}

    \item 假设复合光谱的形成仅与各独立光谱有关。由于复合光谱的荧光信号强度较大，检测灵敏度较高，外界环境对数据造成的干扰可以忽略不计。

    \item 假设实验数据允许存在一定的误差范围。由于测量精度的限制，数据会有一定误差。

    \item 假设叠加光谱均满足瑞利判据。题目中给出的多波峰叠加的复合光谱应当是可以分峰的。

\end{enumerate}