# Level 1: 任选一题（共 10 分）

**Level 1.1【数学】**方程求解：基于 Python，给定函数形式 y=x2-2x+1，给定输入 y，求解 x

**Level 1.2【数学】**数列：基于 Python，给出斐波那契数列在 N 以内的序列，给

定输入 N，求出对应序列。

例如 N=100，即 100 以内的斐波那契数列

**Level 1.3【数学】**数值：基于 Python，用三次函数 y=a\*x^3 拟合出给定 0，1 两点之间的平滑曲线，并作 100 等分，给出相应等分点的值（98+头尾 2 个值）， 给出代码实现

**Level 1.4【物理】**力学分析：一铁球重量为 80 公斤，一开始以 1 米/秒的初速度在水平平面滚动，从开始到静止耗时 10 秒。假设全过程只有摩擦力和重力影响，请基于 Python 给出该球从运动到静止过程的速度-时间图像（时间轴分辨率为 0.1 秒）

**Level 1.5【物理】**力矩分析：给定一个正在做单手俯卧撑的同学，其体重为 70 公斤，脚部触地位置距其单手撑地位置的距离为 1.6m，脚部触地位置距其身体重心位置的距离是 1 米，手掌到肩膀长度为 0.5 米，请基于 Python，给出该同学在肩膀撑起的离地高度 0.5 米过程中的手臂力矩和受力大小

**Level 1.6【物理】**光学：在自动驾驶中，周围物体和车辆的距离和位置（深度信息）是非常重要的信息，通常可以通过不同位置的多个摄像头/相机进行测量。给定两个理想状态的焦距为 f 的相机（无需标定，无畸变），相机像素均为 100 万像素（1000 X 1000 像素），两个相机的光心位置分别是(x1, y)和(x2,

y) (即两个相机水平放置在同一高度)，x1 和 x2 相距 2 米（相机中心距离），假

设某点 P 和相机在同一高度，其在两个相机感光器上的成像点分别为 A 和 B

（假设 AB 距离为 d）。请求出 P 点和两个相机光心连线的垂直距离 D。

提示：根据相机成像原理和相似三角形，可以得出 AB 间距离 d 与相机中心距之比，正比于（D-f）/D

**Level 1.7【统计】**时序统计：COVID-19 是由 SARS-CoV-2 导致的传染性疾病，常见症状包括发烧，咳嗽和呼吸急促。现有全球病情统计结果

（https://github.com/CSSEGISandData/COVID- 19/tree/master/csse\_covid\_19\_data/csse\_covid\_19\_daily\_reports），以数据 04-01- 2022.csv 为例，编写一个程序，输出确诊人数（Confirmed）最多的 10 个国家详细数据（国家 Country/Region，最新更新时间 Last Update，确认病例Confirmed ，死亡病例 Deaths，恢复病例 Recovered）

**Level 1.8【统计】**频率统计：统计《经乱离后天恩流夜郎忆旧游书怀赠江夏韦

太守良宰》（李白，千字诗）的字频，并绘制字频的直方图(Histogram)

提示：https://baike.baidu.com/item/经乱离后天恩流夜郎忆旧游书怀赠江夏韦太守良宰/6865012

**Level 1.9【统计】**数据统计：基于 Python 的 statsmodel, scipy, numpy, pandas, matplotlib 等工具, 根据你的身高增长数据和中国经济总量增长数据（年化身高

和年化 GDP），计算这两个序列的相关性。

提示：https://data.worldbank.org/

**Level 1.10【生化】**分子量计算：给定一个只含 **C，H，O，N** 的化学分子式， 输出它的分子量

**Level 1.11【生化】**Pubchem 是 NCBI 自 2004 年以来开始运营和管理的大型化合物数据库，主要收录了原子数量在 1000 以下以及化学键在 1000 各以下相对较小的分子数据。请基于 PubChemPy 等工具，给出水、苯环和甲醇的 3D 结构式相关信息

**Level 1.12【生化】**请根据 Github 上的 pyscf 等工具，给出水分子 H2O 和氧气分子 O2 的平均场

**Level 1.13【计算机+生化】**基于 Python 及 ncib 提供的 ftp 接口，下载 NCBI 中任意一条/一类指定基因序列，如 bacteria

**Level 1.14【计算机】**基于 Python 的 pygtrans/goslate 等工具，实现中英文文件的双向翻译

**Level 1.15【计算机+统计】**基于 Python 的 pypdf2/pypdf4，对一个 PDF 文件进行文字内容提取

# Level 2: 任选一题（共 10 分）

**Level 2.1【数学】**积分计算：使用 numpy、math 库等工具独立实现二重积分的

计算：

( &



计算

'&

)

𝑥 + 𝑦$𝑑𝑥𝑑𝑦

并与 scipy.integrate 比较二重积分计算效率

提示：统计连续计算 10000 次的耗时

**Level 2.2【数学】**微分方程计算：使用 numpy、sympy、scipy、math 库等工具，根据一阶微分方程 dy/dx=y，画出自变量 x 在定义域（0，5）中对应的 y

值的曲线

提示：该函数为 y=a\*exp(x)，a 为常数

**Level 2.3【数学】**在野外彩弹 CS 游戏中，考虑到重力影响，彩弹的弹道是一个抛物线方程，给定发射点坐标(x0,y0,z0)，落点坐标(x1,y1,z1)，以及彩弹的初速

度 v0（标量数值），使用 numpy、math 库等工具求解满足给定起点、落点和初

速度的彩弹抛物线发射仰角

提示：默认 g=9.8 为常数；如果初速度过小，解可能不存在

**Level 2.4【物理】**给定一个静坐在公交车上的同学，公交车初始速度 v0 已知， 该同学持续记录其手机的角速率  及相对于公交车测量到的当前手机线加速度 a，就可以确定公交车的当前线加速度。以公交车起始速度作为初始条件， 基于以上 v0，序列和 a 序列信息求解公交车的速度（对公交车的线加速度进行积分），以及 t 时刻后相对于初始位置移动的距离（以起始位置作初始条件积分计算）。

**Level 2.5【物理】**西区操场跑道上一跑者体重为 80 公斤，身高为 1.80 米，跑步时身体上下起伏为身高的 10%，任意时刻最多只有一只脚触地，单脚触地时从触地到离地耗时 0.5 秒（触地最低点 0.25 秒，最低点离地 0.25 秒），假设垂直方向上速度变化是分段匀加/减速变化，请基于 Python 给出该跑者从触地到离地过程中脚掌受力变化的受力-时间图像（时间轴分辨率为 0.01 秒）

**Level 2.6【机械+自动化+物理】**在机器人的运动中，不可避免地需要用到电机。在电机运动过程中，电机会接收控制指令进行旋转动作，并感知所旋转到

的实际位置。在实际过程中，我们会发现电机旋转到的实际位置/角度，与想要达到的预期位置/角度总会或多或少存在偏差。目前大多数电机均提供电机动作过程中的精确位置/角度测量，便于闭环控制。请基于 Python 和以上说明，给

出电机转速控制的 PID 算法实现。

提示：机器人、无人机和其他自动化控制里的基础控制方法雏形

**Level 2.7【统计+金融】**统计检验：基于 Python 的 scipy, statsmodel, numpy, pandas\_datareader 等工具, 获取美的集团（000333）股票在最近三个月内的收盘价数据，对其进行游程检验（RUNS test）并给出检验结果

**Level 2.8【统计】**接 Level 1.1【统计】，基于 Python 的 scipy, statsmodel, numpy, pandas 等工具，根据新加坡 2021 年 10 月到 12 月共三个月的 COVID 疫情数据，对确诊病例数量进行 logistic 回归分析并建模，并给出 2022 年 1 月的

预测曲线

提示：https://github.com/CSSEGISandData

**Level 2.9【统计+经济】**基于世界银行数据，下载世界各国历年的 GDP 总值和增长率数据，基于 Python 的 scipy, statsmodel, numpy, pandas\_datareader 等工

具，提取金砖四国（巴西、俄罗斯、印度、中国）近 20 年的 GDP 总值和增长率数据，求出这 4 个国家 GDP 总值和增长率的斯皮尔曼相关系数和肯德尔相

关系数

提示 1：https://data.worldbank.org/

提示 2：可使用 Pandas 中的 Series.corr()

**Level 2.10【生化】**基于 Python 中 Bio, numpy, pandas 等工具来比较两个基因序

列’ TTTGGAAAAC’和’ TTTTTTACCG’的相似性

提示：当前疫情病毒变异树的雏形，通过比较和分析不同变种的相似性而得到

**Level 2.11.【生化】**DNA 转录 RNA 过程模拟：在 RNA 转录过程中，一条DNA 链用作模板，mRNA 通过一次复制一个核苷酸来构建 RNA 链，转录过程中使用尿嘧啶代替胸腺嘧啶。使用字符代表碱基，DNA 字符串含有 A，C， G，T，RNA 字符串含有 A，C，G，U。现有 DNA 字符串 s，将碱基 T 替换为碱基 U，求转录后的 RNA。如 DNA（ATGGAACTTGACTACGTAAATT）  RNA（GAUGGAACUUGACUACGUAAAUU）

**Level 2.12【生化+计算机】**生物显微镜照片处理过程中，往往会将一张彩色图像（RGB 三色）的三种颜色通道分别处理为三张灰度图像（每张灰度图像对应一个颜色通道），图像中每个像素的灰度值大小表示该颜色的明暗程度。但 是，因为显微镜进光量的不同，同一标本不同通道的明暗程度会有较大不同

（例如红色通道偏亮，绿色通道偏暗）。请基于 Python，调整一张彩色图片的

不同通道的明暗程度，使得图片不同通道的明暗程度趋于一致。

提示：对不同通道的灰度图像中每个像素的灰度值都加/减去一个不同的值，类似标准化的处理

提示：天文/生化显微镜数据预处理的常见操作

**Level 2.13【计算机】**基于 Python，使用浏览器引擎 Selenium 爬取科大首页中的科大要闻第一条新闻，要求爬取结果至少有新闻标题、时间、以及新闻内容**Level 2.14【计算机】**对手机拍摄的任意照片进行 OCR 识别，例如校园内的咖啡厅付款小票、车牌、路标、门牌等带有文字和数字的照片，基于现有的 OCR

库，对多类（至少三类）照片进行文字和数字提取

提示：python 提供多种 OCR 库

**Level 2.15【计算机】**基于 Python 的 librosa 及其 effects 子库等工具，对自己的

语音录制文件进行音调（pitch）的修改，例如中音改到高音

提示：对声音进行任意操作，类似导航系统中的个性化语音

# Level 3: 任选一题（共 10 分）

**Level 3.1【数学】**扩展 Python 的 Num 类型，使其支持四元数的计算，并实现欧拉角和四元数的相互转换计算

**Level 3.2 【数学】**微分计算：使用 numpy、sympy、scipy、math 库等工具求解二阶常系数齐次微分方程 a\*f(x) + b\*f ’(x) + c\*f ’’(x) = 0

**Level 3.3【数学】**数值：接 Level 1.3【数学】，基于 Python，给出指定序列的三次样条插值方法的代码实现，例如给定输入序列 y【0，0.5，0.8，1，0.4，

0】和 x【0，1，2，3，4，5】，求出对应插值点（本段序列给出 100 个插值点

即可）。

提示 1：三次样条插值是把已知数据分割成若干段，每段构造一个三次函数， 使得分段函数的衔接处具有 0 阶连续，一阶导数连续，二阶导数连续的性质提示 2：机器人控制里常用的加速度调节方法

**Level 3.4（加 1 分）【数学】**方程：接物理 Level 1.3， 在自动驾驶中，周围物体和车辆的距离和位置（深度信息）是非常重要的信息，通常可以通过不同位置的多个摄像头/相机进行测量。物理 Level 1.3 给出了双目测距方案，在多目测距中，可以按照每一对相机列出对应的方程，进而形成方程组，用方程组的方式进行求解。但由于系统误差、测量误差以及偏差等因素存在，所求得的位置并不精确，牛顿法和高斯牛顿法表现不佳，请基于 Python，给出列文伯格-马夸

尔特（Levenberg-Marquardt）算法的代码实现

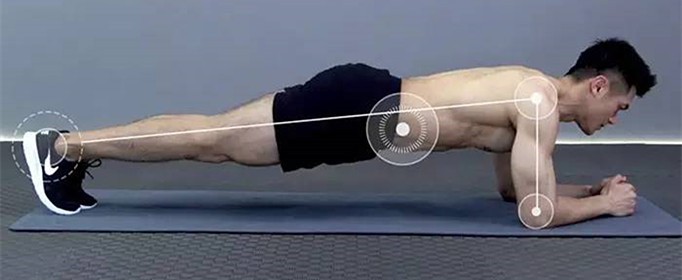
提示 1：列-马 LM 算法是使用最广泛的非线性最小二乘算法，其利用梯度求最大（小）值的算法，同时具有梯度法和牛顿法的优点

提示 2：自动驾驶、动作捕捉等的多目测距中常用计算方法

**Level 3.5【物理+数学+计算机】**给定正在做平板支撑的一位同学，基于以下图

像请给出其身体姿态的力矩求解，并计算其腹部重心位置肌肉的受力情况。

提示：可基于图像识别出身体重心和手脚的像素位置，假设身高、体重已知



**Level 3.6（加 1 分）【物理】**- 接【数学】Level 2.3 野外彩弹 CS 游戏，考虑重力和空气阻力影响，给定发射点坐标(x0,y0,z0)，落点坐标(x1,y1,z1)，空气阻力f（矢量，以(fx,fy,fz)表示），及彩弹的初速度大小 v0（标量数值），使用

numpy、math 库等工具求解满足给定起点、落点、空气阻力和初速的发射仰角

提示 1：默认 g=9.8 为常数；如果初速度过小，解可能不存在提示 2：弹道计算机的雏形

**Level 3.7（加 1 分）【自动化+数学+物理】**接 Level 2.1【物理】机器人和现代驾驶系统中的导航系统的传感器数据（加速度、陀螺仪）普遍存在小误差，会随时间累积成大误差，其误差大体上与时间成正比，因此需要不断进行修正。而全球定位系统 GPS 的定位精度是 1 到 30 米，长距离来看误差较小，但短时间和短距离内误差较大。现代惯性导航系统使用各种信号（例如全球定位系统GPS 等）对其进行修正，采取控制论方法如 EKF 扩展卡曼滤波等对不同信号进行加权过滤，从而保证惯性导航系统的精度及可靠性。

请基于加速度、陀螺仪、磁力计和 GPS 数据，基于 Python 和四元数，给出扩

展卡曼滤波算法（EKF）的代码实现

提示：捷联惯导系统、无人机等的位姿获取方法雏形

**Level 3.8【统计+金融】**协整检验：基于 Python 获取包括美的集团（000333） 和贵州茅台（600519）的沪深 300 指数中的所有 300 只股票近 1 年的每日调整

后收盘价（adjusted close）数据，基于 Python 的 statsmodel, numpy 等工具, 对过去 1 年美的集团（000333）和贵州茅台（600519）股票进行协整检验（co- integration test）并给出检验结果，并进一步给出这 300 只中协整关系表现最好

的 10 对股票。

提示 1：使用 statsmodels 中的 statsmodels.tsa.stattools.coint 提示 2：建议考虑多进程/多线程实现

提示 3：这些股票对，有什么价值？

**Level 3.9【统计+生化+天文+计算机】**生物显微镜/天文望远镜的图像往往用于记录描述不同物质/细胞/病毒，这些目标物质/细胞/病毒一般会具有显著区别于周边区域的亮度（亮斑），若同一标本的某个通道灰度图像在某一位置/区域出现了显著区别于周边区域的亮度，则需要对其进行标记并计数，请给出该通道

图像的这些亮斑。

提示 1：计算出灰度图像的基本底色亮度，找到灰度图像中显著亮于该背景灰度的所有像素，对这些像素，相邻像素做聚合，则得到亮斑，对其计数即可 提示 2：：高像素图像可考虑切割成多个边缘略有重叠的区块后，用多进程/多线程实现

**Level 3.10【统计+天文+计算机】**基于 SDSS 数据库的 skyserver 中近一个月内的巡天图，请基于 Python 进行图像识别，给出巡天图中天狼星 A 所在像素位置，并给出该星在近一个月内的亮度变化曲线及相关统计量（至少三个统计

量）。

提示 1：该星应为北天区最亮的星（可按明暗程度快速查找候选像素）

提示 2：如果识别出多颗星及其像素位置，还可以进一步确定照片在地球上的拍摄地点，即星光导航的雏形

提示 3：对恒星/行星的星光强度（行星为反射光）进行持续追踪并进行频域变换，还可以进一步得到这些星的运行周期，可以快速发小论文

**Level 3.11【生化+数学】**化学式配平

输入化学方程式，输出配方之后的结果（公式不包含除字母、下表以及+= 之外的符号）

例如输入： KMnO4+H2S+H2SO4=MnSO4+K2SO4+S+H2O

输出：2KMnO4+5H2S+3H2SO4=2MnSO4+K2SO4+5S+8H2O

**Level 3.12【计算机+数学+物理+自控】**给定自己手机（含加速度、陀螺仪、磁力计）的初始位置和姿态，基于其传感器数据序列，利用四元数给出该设备在运动过程中的连续姿态变化

提示：重点是从手机获取，打通手机的接口，可考虑 android 手机

**Level 3.13【计算机+数学+物理】**对手机图像进行修改，给出三种自定义的光线

滤镜

提示：当前手机/app 滤镜效果的雏形

**Level 3.14（加 1 分）【计算机】**基于 Python 抓取自己手机微信所关注的某个公众号中含有指定关键词的文章，爬取结果包括文章标题、发布时间、文章内

容、打赏（喜欢）人数、阅读数、点赞数和在看数

提示：可基于浏览器引擎 Selenium 或 android 虚拟机

**Level 3.15（加 1 分）【计算机】**基于 Python 的 faceswap 等工具，对自己的视

频文件进行人脸替换，例如改为动画人物

提示：类似魔术/川剧中的变脸