2021阿里巴巴全球数学竞赛-预选赛

1. 在一个虚拟的世界中, 每个居民(设想为没有大小的几何点) 依次编号为 $1, 2, \cdots$. 为了抗击某种疫情, 这些居民要接种某疫苗，并在注射后在现场留观一段时间。现在假设留观的场所是平面上的一个半径为 $\frac{1}{4}$ 的圆周。为了安全，要求第 m 号居民和第 n 号居民之间的距离 $d_{m,n}$ 满足

$$(m+n)d_{m,n} \geq 1,$$

这里我们考虑的是圆周上的距离, 也就是两点间劣弧的弧长。那么

(i) 下列选项( ) 符合实际情况。

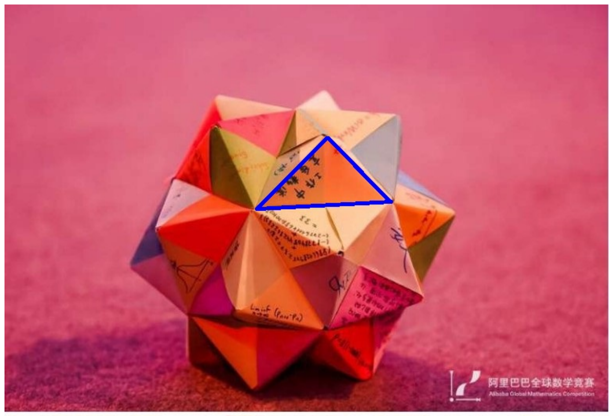
A 这个留观室最多能容纳 8 个居民;

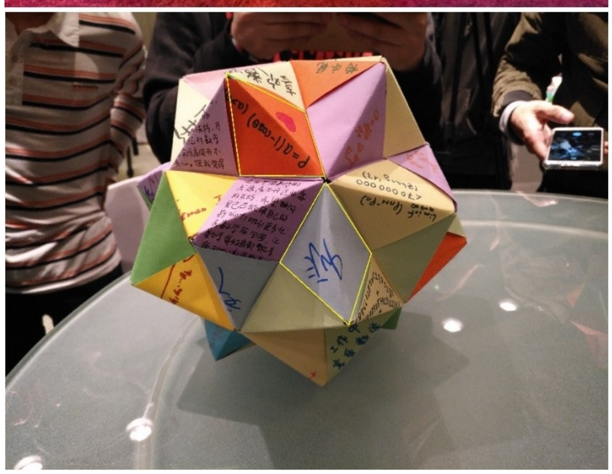
B 这个留观室能容纳的居民个数有大于 8 的上限;

C 这个留观室可以容纳任意多个居民。

(ii) 证明你的论断。

2. 2019年第一届阿里巴巴数学竞赛的优胜者们在参加集训营的时候, 集体送给主办方负责人的礼物, 是一个有 60 个全等的三角形面的多面体。从图中我们可以看到, 这个多面体的表面是 60 个全等的空间四边形拼接而成的。





一个空间 n 边形是指由一个平面 n 边形沿若干条对角线做适当翻折(即在选定的对角线处形成适当的二面角) 后得到的空间图形。两个空间图形全等指的是它们可以通过 $\mathbb{R}^3$ 中的一个等距变换完全重合。一个多面体指的是一个空间有界区域, 其边界可以由有限多个平面多边形沿公共边拼接而成。

(i) 我们知道 $2021 = 43 \times 47$. 那么是否存在一个多面体, 它的表面可以由 43 个全等的空间 47 边形拼接而成?

(ii) 请对你的判断给出逻辑的解释。

3. 去年，张师傅因为多旋圈面爆红，今年他来到了达摩院给扫地僧做面。某天，软件工程师小李跟张师傅吐槽工作。小李主要研究和设计算法用于调节各种产品的参数。这样的参数一般可以通过极小化 $\mathbb{R}^n$ 上的某个损失函数 f 求得。在小李最近的一个项目中，这个损失函数是另外一个课题组提供的；出于安全考虑和技术原因，该课题组难以向小李给出此函数的内部细节，而只能提供一个接口用于计算任意 $\mathbf{x} \in \mathbb{R}^n$ 处的函数值 $f(\mathbf{x})$。所以，小李必须仅基于函数值来极小化 f。而且，每次计算 f 的值都消耗不小的计算资源。好在该问题的维度 n 不是很高 (10 左右)。另外，提供函数的同事还告知小李不妨先假设 f 是光滑的。

这个问题让张师傅想起了自己收藏的一台古董收音机。要在这台收音机上收听一个节目，你需要小心地来回拧一个调频旋钮，同时注意收音效果，直到达到最佳。在这过程中，没有人确切地知道旋钮的角度和收音效果之间的定量关系是什么。张师傅和小李意识到，极小化 f 不过就是调节一台有多个旋钮的机器: 想象 $\mathbf{x}$ 的每一个分量由一个旋钮控制，而 $f(\mathbf{x})$ 表示这台机器的某种性能，只要我们来回调整每个旋钮，同时监视 f 的值，应该就有希望找到最佳的 $\mathbf{x}$。受此启发，两人一起提出了极小化 f 的一个迭代算法，并命名为"自动前后调整算法" (Automated Forward/Backward Tuning，AFBT，算法 1)。在第 k 次迭代中，AFBT 通过前后调整 $\mathbf{x}_k$ 的单个分量得到 $2n$ 个点 $\{\mathbf{x}_k \pm t_k\mathbf{e}^i : i = 1, \ldots, n\}$, 其中 t_k 为步长；然后，令 $\mathbf{y}_k$ 为这些点中函数值最小的一个，并检查 $\mathbf{y}_k$ 是否使 f 充分减小；若是，取 $\mathbf{x}_{k+1} = \mathbf{y}_k$，并将步长增倍；否则，令 $\mathbf{x}_{k+1} = \mathbf{x}_k$ 并将步长减半。在算法 1 中， $\mathbf{e}^i$ 表示 $\mathbb{R}^n$ 中的第 i 个坐标向量，它的第 i 个分量为 1, 其余皆为 0；$\mathbb{1}(\cdot)$ 为指示函数 —— 若 $f(\mathbf{x}_k) - f(\mathbf{y}_k)$ 至少为 t_k 之平方，则 $\mathbb{1}[f(\mathbf{x}_k) - f(\mathbf{y}_k) \geq t_k^2]$ 取值为 1，否则为 0。

算法 1 自动前后调整算法 (AFBT)

输入 $\mathbf{x}_0 \in \mathbb{R}^n$, $t_0 > 0$。对 $k = 0, 1, 2, \ldots,$ 执行以下循环。

1: $\mathbf{y}_k := \operatorname{argmin}\{f(\mathbf{y}) : \mathbf{y} = \mathbf{x}_k \pm t_k\mathbf{e}^i,\ i = 1, \ldots, n\}$ # 计算损失函数。
2: $s_k := \mathbb{1}[f(\mathbf{x}_k) - f(\mathbf{y}_k) \geq t_k^2]$ # 是否充分下降？ 是：$s_k = 1$； 否：$s_k = 0$。
3: $\mathbf{x}_{k+1} := (1 - s_k)\mathbf{x}_k + s_k\mathbf{y}_k$ # 更新迭代点。
4: $t_{k+1} := 2^{2s_k-1}t_k$ # 更新步长。$s_k = 1$：步长增倍；$s_k = 0$：步长减半。

现在，我们对损失函数 $f : \mathbb{R}^n \to \mathbb{R}$ 作出如下假设。

假设 1. f 为凸函数，即对任何 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ 与 $\alpha \in [0, 1]$ 都有

$$f((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) \ \leq \ (1-\alpha)f(\mathbf{x}) + \alpha f(\mathbf{y}).$$

假设 2. f 在 $\mathbb{R}^n$ 上可微且 ∇f 在 $\mathbb{R}^n$ 上 L-Lipschitz 连续。

假设 3. f 的水平集有界，即对任意 $\lambda \in \mathbb{R}$，集合$\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq \lambda\}$ 皆有界。

基于**假设 1** 与**假设 2**，可以证明

$$\langle \nabla f(\mathbf{x}),\ \mathbf{y}-\mathbf{x}\rangle \ \le\ f(\mathbf{y})-f(\mathbf{x}) \ \le\ \langle \nabla f(\mathbf{x}),\ \mathbf{y}-\mathbf{x}\rangle + \frac{L}{2}\|\mathbf{x}-\mathbf{y}\|^2$$

对任何 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ 成立；**假设 1** 与**假设 3** 则保证f 在 $\mathbb{R}^n$ 上取到有限的最小值 f^*。凸函数的更多性质可参考任何一本凸分析教科书。

在**假设 1–3**下，对于 AFBT，证明

$$\lim_{k\to\infty} f(\mathbf{x}_k) \ =\ f^*.$$

4. 令 n 为正整数。对任一正整数 k，记 $0_k = \mathrm{diag}\{\underbrace{0, \ldots, 0}_{k}\}$ 为 $k \times k$ 的零矩阵。令

$$Y = \begin{pmatrix} 0_n & A \\ A^t & 0_{n+1} \end{pmatrix}$$

为一个 $(2n+1) \times (2n+1)$ 矩阵, 其中 $A = (x_{i,j})_{1 \le i \le n, 1 \le j \le n+1}$ 是一个 $n \times (n+1)$ 实矩阵且 A^t 记 A 的转置矩阵，即 $(n+1) \times n$ 的矩阵，(j, i) 处元素为 $x_{i,j}$.

(i) 称复数 λ 为 $k \times k$ 矩阵 X 的一个特征值, 如果存在非零列向量

$$v = (x_1, \ldots, x_k)^t$$

使得 $Xv = \lambda v$. 证明: 0 是 Y 的特征值且 Y 的其他特征值形如 $\pm\sqrt{\lambda}$, 其中非负实数 λ 是 AA^t 的特征值。

(ii) 令 $n = 3$ 且 a_1, a_2, a_3, a_4 是 4 个互不相等的正实数。记

$$a = \sqrt{\sum_{1 \le i \le 4} a_i^2}$$

以及

$$x_{i,j} = a_i \delta_{i,j} + a_j \delta_{4,j} - \frac{1}{a^2}(a_i^2 + a_4^2) a_j$$

$(1 \le i \le 3,\ 1 \le j \le 4)$, 其中 $\delta_{i,j} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \ne j \end{cases}$. 证明: Y 有 7 个互不相等的特征值。

5. 对于 $\mathbb{R}$ 上的连续且绝对可积的复数值函数 $f(x)$，定义 $\mathbb{R}$ 上的函数 $(Sf)(x)$:

$$(Sf)(x)=\int_{-\infty}^{+\infty} e^{2\pi \mathbf{i} ux} f(u)\mathrm{d}u.$$

(i) 求 $S(\frac{1}{1+x^2})$ 和 $S(\frac{1}{(1+x^2)^2})$ 的显式表达式。

(ii) 对任意整数 k，记 $f_k(x)=(1+x^2)^{-1-k}$. 假设 $k\geq 1$，找到常数 c_1, c_2 使得函数 $y=(Sf_k)(x)$ 满足二阶常微分方程

$$xy''+c_1y'+c_2xy=0.$$

6. 当某公司推出一个新的社交软件时，公司的市场部门除了会关心该软件的活跃客户的总人数随时间的变化，也会对客户群体的一些特征做具体的调研和分析。我们用 $n(t,x)$ 表示客户的数量密度（以下简称密度），这里 t 表示时间，而 x 表示客户对该社交软件的使用时长，那么在 t 时刻，对于 $0<x_1<x_2$，使用时长介于 x_1 和 x_2 之间的客户数量为 $\int_{x_1}^{x_2} n(t,x)dx$。我们假设，密度 $n(t,x)$ 随着时间演化受以下几个因素的影响：

假设 1. 当客户持续使用该社交软件时，他的使用时长随时间线性增长。

假设 2. 客户在使用过程中，可能会停止使用，我们假设停止速率 $d(x)>0$ 只跟使用时长 x 有关。

假设 3. 新客户的来源有两个。

① 公司的宣传：单位时间内因此增加的人数是时间的函数，用 $c(t)$ 表示。

② 老客户的宣传：老客户会主动向自己的同事、朋友等推荐使用该社交软件，推荐成功的速率跟客户的使用时长 x 有关，记作 $b(x)$。

假设如果在某一时刻, 记为 $t=0$ 时，密度函数是已知的，$n(0,x)=n_0(x)$。可以推导出，$n(t,x)$ 的时间演化满足如下的方程

$$\begin{cases} \frac{\partial}{\partial t}n(t,x)+\frac{\partial}{\partial x}n(t,x)+d(x)n(t,x)=0, \quad t\geq 0, x\geq 0, \\ N(t):=n(t,x=0)=c(t)+\int_0^\infty b(y)n(t,y)dy. \end{cases} \tag{1}$$

这里 $N(t)$ 可解读为新客户的增加速率。我们假设 $b,d\in L_+^\infty(0,\infty)$，即 $b(x)$ 和 $d(x)$ 正且(本质)有界。以下，我们先做一个简化假设： $c(t)\equiv 0$，即新客户的增加只跟老客户的宣传有关。

(i) 根据**假设 1**和**假设 2**，形式地推导出(1) 中 $n(t,x)$ 所满足的偏微分方程，需要在推导过程中指出模型假设和数学表达式之间的对应关系。再根据**假设 3**，解释(1) 中 $N(t)$ 的定义的含义。

(ii) 我们想要研究新客户的增加速率 $N(t)$ 和推荐成功速率 $b(x)$ 之间的关系。为此，请推导出一个 $N(t)$ 所满足的方程，且方程中只包含 $N(t)$, $n_0(x)$, $b(x)$, $d(x)$，而不包含 $n(t,x)$。并证明， $N(t)$ 满足如下估计

$$|N(t)|\leq \|b\|_\infty e^{\|b\|_\infty t}\int_0^\infty |n_0(x)|\,dx, \tag{2}$$

这里 $\|\cdot\|_\infty$ 表示 L^∞ 范数。

(iii) 最后，我们想要研究，在充分长的时间之后，数量密度函数 $n(t,x)$ 有什么渐近的趋势。由于客户总人数可能一直在增加，所以我们不方便直接研究数量密度函数 $n(t,x)$，而更应该去看一个重整化的的密度函数。

为此，我们首先假设如下的特征值问题有唯一解 $(\lambda_0, \varphi(x))$：

$$\begin{cases} \varphi'(x) + (\lambda_0 + d(x))\,\varphi(x) = 0, \quad x \geq 0, \\ \varphi(x) > 0, \quad \varphi(0) = \int_0^\infty b(x)\varphi(x)dx = 1, \end{cases}$$

并且它的对偶问题也有唯一的解 $\psi(x)$:

$$\begin{cases} -\psi'(x) + (\lambda_0 + d(x))\,\psi(x) = \psi(0)b(x), \quad x \geq 0, \\ \psi(x) \geq 0, \quad \int_0^\infty \psi(x)\varphi(x)dx = 1. \end{cases}$$

然后，我们定义重整化密度 $\tilde{n}(t,x) := n(t,x)e^{-\lambda_0 t}$。证明，对于任意凸函数 $H : \mathbb{R}^+ \to \mathbb{R}^+$ 满足 $H(0) = 0$, 我们有

$$\frac{d}{dt}\int_0^\infty \psi(x)\varphi(x)H\left(\frac{\tilde{n}(t,x)}{\varphi(x)}\right)dx \leq 0, \quad \forall t \geq 0,$$

并证明

$$\int_0^\infty \psi(x)n(t,x)dx = e^{\lambda_0 t}\int_0^\infty \psi(x)n_0(x)dx.$$

为了简化证明，我们在演算中假定在 ∞ 处的边界项的贡献都是可以忽略的。