阿里巴巴全球数学竞赛决赛题目集锦

代数与数论

问题1

令p 为奇素数, $m \geq 0$和$N \geq 1$ 为整数。设Λ 为一秩为$2m+1$ 的自由$\mathbb{Z}/p^N\mathbb{Z}$-模，并设

$$(\ ,\ ):\ \Lambda \times \Lambda \to \mathbb{Z}/p^N\mathbb{Z}$$

为一完美对称$\mathbb{Z}/p^N\mathbb{Z}$-双线性型。这里, “完美” 指的是诱导映射

$$\Lambda \to \mathrm{Hom}_{\mathbb{Z}/p^N\mathbb{Z}}(\Lambda, \mathbb{Z}/p^N\mathbb{Z}), \quad x \mapsto (x, \cdot)$$

为同构。试求集合

$$\{x \in \Lambda \mid (x, x) = 0\}$$

的元素个数，表达为p, m, N 的函数。

答案1

对每个整数$0 \leq n \leq N$, 设

$$\Lambda(n) \coloneqq \{x \in \Lambda \mid (x, x) \in p^n\mathbb{Z}/p^N\mathbb{Z}\},$$

它在$p^n\Lambda$ 的迁移下是稳定的. 设$C(n) \coloneqq |\Lambda(n)|$. 我们的目标是计算$C(N)$. $C(0) = |\Lambda| = p^{(2m+1)N}$ 是平凡的.

我们首先算一下$C(1)$. 商空间$\Lambda/p\Lambda$, 关于$\mathbb{F}_p$-双线性形式$(\ ,\ ) \bmod p$, 是$\mathbb{F}_p$ 上的一个$2m+1$ 维非退化的二次空间. 显然一个元素$x \in \Lambda/p\Lambda$ 属于$\Lambda(1)/p\Lambda$ 当且仅当x 的长度为0.

我们断言对每个$2m+1$ 维$\mathbb{F}_p$ 系数的非退化二次空间V, 长度为0 的元素个数是p^{2m}. 下面对m 作归纳. $m = 0$ 时是平凡的. $m > 0$ 时, 取$v_1, v_2 \in V$ 使得$(v_1, v_1) = 1$, $(v_2, v_2) = -1$, 且$(v_1, v_2) = 0$. 设V' 是$\{v_1, v_2\}$ 在V 中的正交补空间, 它是一个$2m-1$ 维非退化的二次空间. 关于该正交分解, 每个$x \in V$ 可以写成$x = x_1v_1 + x_2v_2 + x'$, 其中$x_1, x_2 \in \mathbb{F}_p$, $x' \in V'$. 于是$(x, x) = x_1^2 - x_2^2 + (x', x')$. 为了计算满足$(x, x) = 0$ 的x 的元素个数, 考虑以下两种情况: 若$(x', x') = 0$, 则满足$x_1^2 - x_2^2 = 0$ 的$(x_1, x_2) \in \mathbb{F}_p^2$ 的个数是$2p-1$. 若$(x', x') \neq 0$, 则满足$x_1^2 - x_2^2 = -(x', x') \neq 0$ 的$(x_1, x_2) \in \mathbb{F}_p^2$ 的个数是$p-1$ (因$p \neq 2$). 从而由归纳假设知, 长度为0 的x 的元素个数共有$p^{2m-2}(2p-1) + (p^{2m-1} - p^{2m-2})(p-1) = p^{2m}$ 个.

总之, $C(1) = p^{2m}p^{(2m+1)(N-1)}$.

一般地, 对$n \geq 1$, 设$\Lambda(n)' \coloneqq \Lambda(n) \setminus p\Lambda$ 以及$\Lambda(n)'' \coloneqq \Lambda(n) \cap p\Lambda$, 定义$C(n)'$ 与$C(n)''$ 分别为集合$\Lambda(n)'$ 与$\Lambda(n)''$ 的元素个数. 我们断言

1. $n \geq 2$ 时, $\Lambda(n-2)/p^{n-1}\Lambda \to \Lambda(n)''/p^n\Lambda$ 的乘p 映射是双射.
2. $n \geq 2$ 时, 由包含映射$\Lambda(n)' \to \Lambda(n-1)'$ 诱导的映射$\Lambda(n)'/p^n\Lambda \to \Lambda(n-1)'/p^{n-1}\Lambda$ 是一个p^{2m}-对-1 的映射.

这两个断言的证明稍后给出, 我们首先由此计算$C(n)$. 由第一个断言可知$C(n)'' = p^{-(2m+1)}C(n-2)$, $n \geq 2$; 而由第二个断言知$C(n)' = p^{-1}C(n-1)'$, $n \geq 2$. 对于后者, 我们有

$$C(n)' = p^{-(n-1)}C(1)' = p^{-(n-1)}(C(1) - |p\Lambda|) = p^{(2m+1)(N-1)-(n-1)}(p^{2m} - 1).$$

结合前者可得以下的递推公式

$$C(n) = C(n)' + C(n)'' = p^{-(2m+1)}C(n-2) + p^{(2m+1)(N-1)-(n-1)}(p^{2m} - 1).$$

进一步容易得到

$$C(N) = p^{(2m+1)r+2m(N-2r)} + \frac{p^{(2m-1)r} - 1}{p^{(2m-1)} - 1}p^{(2m+1)r-1+2m(N-2r)}(p^{2m} - 1)$$

其中$r := \lfloor N/2 \rfloor$.

接下来证明以上的两个断言.

前者可以通过以下的观察得到, 一个元素$x \in \Lambda$ 属于$\Lambda(n-2)$ 当且仅当px 属于$\Lambda(n)$.

对于后者, 取定$y \in \Lambda(n-1)'/p^{n-1}\Lambda$. 其原像由以下元素组成: $y + p^{n-1}x$ 其中$x \in \Lambda/p\Lambda$ 且

$$(y, y) + 2p^{n-1}(x, y) + p^{2n-2}(x, x) \in p^n\mathbb{Z}/p^N\mathbb{Z}.$$

因$n \geq 2$ 以及$(y, y) \in p^{n-1}\mathbb{Z}/p^N\mathbb{Z}$, 上式等价于

$$p^{1-n}(y, y) + 2(x, y) \equiv 0 \mod p.$$

由于$p \neq 2$, $y \notin p\Lambda/p^{n-1}\Lambda$, 以及$\Lambda/p\Lambda$ 是一个非退化的二次空间, 上式是一个$\Lambda/p\Lambda$ 上关于x 的一次多项式, 从而它的解集的元素个数是p^{2m}.

问题2

设 p 是一个大于 2023 的素数。记 $\mathcal{X}$ 为 $\mathbb{F}_p$-线性空间 $\mathbb{F}_p^{2023}$ 的所有 2000 维子空间组成的集合。找出满足对任意 $V \in \mathcal{X}$，

$$\sum_{W \in \mathcal{Y}} (V \cap W) = V$$

成立的 $\mathcal{X}$ 的子集 $\mathcal{Y}$ 的最小可能元素个数；并证明你的答案。

答案2

最小元素个数是25.

我们首先证明存在$\mathcal{Y} \subset \mathcal{X}$, 其元素个数是25 并且满足对每个$V \in \mathcal{X}$, 有

$$\sum_{W \in \mathcal{Y}} (V \cap W) = V. \tag{1}$$

设$\mathbb{F}_p^{2023}$ 的标准基是$\{e_1,\ldots,e_{2023}\}$. 对$1\leqslant k\leqslant 25$, 令

$$W_k \coloneqq \operatorname{span}\left\{\{e_1,\ldots,e_{2023}\}\setminus\{e_{23(k-1)+1},\ldots,e_{23k}\}\right\}.$$

我们证明$\mathcal{Y}\coloneqq\{W_1,\ldots,W_{25}\}$ 满足要求. 取$V\in\mathcal{X}$ 并注意到$\sum_{W\in\mathcal{Y}}(V\cap W)\subset V$. 若(1) 对$V$ 不成立, 则

$$d\coloneqq \dim\sum_{W\in\mathcal{Y}}(V\cap W)\leqslant 1999.$$

这种情形下, 对每个$W\in\mathcal{Y}$, $V\cap W$ 在$\sum_{W\in\mathcal{Y}}(V\cap W)$ 中的余维数至多是$d-(2000-23)\leqslant 22$. 于是

$$\dim\bigcap_{W\in\mathcal{Y}}(V\cap W)\geqslant 1999-25\cdot 22=1449.$$

另一方面, 有

$$\dim\bigcap_{W\in\mathcal{Y}}(V\cap W)\leqslant\dim\bigcap_{W\in\mathcal{Y}}W=2023-25\cdot 23=1448,$$

两者矛盾.

以下只需证明对每个元素个数至多是24 的子集$\mathcal{Y}\subset\mathcal{X}$, 存在$V\in\mathcal{X}$, 使(1) 不成立. 记$(\mathbb{F}_p^{2023})^\vee$ 为$\mathbb{F}_p^{2023}$ 的$\mathbb{F}_p$-线性对偶. 对每个$W\in\mathcal{Y}$, 取非零元素$f_W\in(\mathbb{F}_p^{2023})^\vee$ 使得$W\subset\operatorname{Ker}f_W$. 设$M$ 是一个包含$\{f_W\,|\,W\in\mathcal{Y}\}$ 的$(\mathbb{F}_p^{2023})^\vee$ 的一个24 维子空间. 因$p>24$, 则存在M 的一个23 维子空间N 满足$N\cap\{f_W\,|\,W\in\mathcal{Y}\}=\emptyset$. 令$V\coloneqq\bigcap_{f\in N}\operatorname{Ker}f$. 则$V$ 属于$\mathcal{X}$ 且满足对每个$W\in\mathcal{Y}$, $V\cap\operatorname{Ker}f_W=\bigcap_{f\in M}\operatorname{Ker}f$. 于是

$$\sum_{W\in\mathcal{Y}}(V\cap W)\subset\sum_{W\in\mathcal{Y}}(V\cap\operatorname{Ker}f_W)=\bigcap_{f\in M}\operatorname{Ker}f,$$

从而真包含于V.

几何与拓扑

问题1

当$n \geq 2$ 时, 记$M_n = \{(u,v) \in S^n \times S^n | u \cdot v = 0\}$, 其中$u \cdot v$ 是u, v 之间的欧氏内积. 假定M_n 的拓扑从$S^n \times S^n$ 诱导而来. 证明:

(1) M_n 为$S^n \times S^n$ 的连通正则子流形;

(2) M_n 为Lie 群当且仅当$n = 2$.

答案1

(1) 考虑函数$f : S^n \times S^n \to \mathbb{R}, f(u,v) = u \cdot v$. 如果$u_0 \neq \pm v_0$, 我们证明$f$ 在(u_0, v_0) 处的秩是1. 事实上，令$w_0' = v_0 - (u_0 \cdot v_0)u_0$, 则

$$w_0' \cdot u_0 = 0, w_0' \cdot v_0 = 1 - (u_0 \cdot v_0)^2 > 0.$$

设$w_0 = w_0'/\|w_0'\| \in S^n$, 定义

$$\sigma(t) = (u_0 \cos t + w_0 \sin t, v_0), t \in \mathbb{R}.$$

则σ 是$S^n \times S^n$ 上从(u_0, v_0) 出发的一条曲线. 由于

$$f_{*(u_0,v_0)}\sigma'(0) = \frac{d}{dt}\bigg|_{t=0} f(\sigma(t)) = w_0 \cdot v_0 > 0,$$

f 在(u_0, v_0) 处的秩是1. 特别地, f 在$M_n = f^{-1}(0)$ 处的秩同样是1, 因此M_n 是$S^n \times S^n$ 的一个正则子流形. 作为微分流形, M_n 恰好是S_n 的切丛中的单位球面丛，从而是连通的。

(2) $n = 2$ 时, 考虑映射

$$\phi : M_2 \to SO(3), (u,v) \mapsto (u, v, u \times v).$$

ϕ 是一个微分同胚, 从而诱导了M_2 上的Lie 群结构.

$n = 3$ 时, 由于S^3 的切丛是平凡丛, $M_3 = S^3 \times S^2$. 特别地, M_3 是单连通的并且$H^2(M_3; \mathbb{R})$ 是非平凡的, 从而根据标准的事实, M_3 不是一个Lie 群.

$n = 4$ 时, 考虑Gysin 序列

$$\to H^{i-4}(S^4;\mathbb{R}) \overset{\cup e}{\to} H^i(S^4;\mathbb{R}) \to H^i(M^4;\mathbb{R}) \to H^{i-3}(S^4;\mathbb{R}) \overset{\cup e}{\to} H^{i+1}(S^4;\mathbb{R}) \to \cdots$$

因TS^4 的欧拉类是非凡的, 于是$H^3(M_4; \mathbb{R})$ 是平凡的. 同样地, 由于M_4 是单连通的, 由标准的事实知它不是一个Lie 群.

$n \geq 5$时, 由Gysin 序列不难看出$H^3(M_n; \mathbb{R})$ 是平凡的. 与上面同理可知M_n 不是一个Lie 群.

注: 任何单连通的紧Lie 群的H^2 是平凡的且H^3 是非平凡的.

问题2

设M 是维数黎曼轨形(orbifold)，维数$n \geq 2$，具有孤立奇点。

(1) 假设M 的截面曲率非负，无穷远处的体积增长大于0，即，对某个$x \in M$，

$$\lim_{r\to\infty} \frac{vol(B(x,r))}{r^n} > 0. \tag{2}$$

证明：M 的奇异点最多只有1 个。

(2) 假设M 紧致，里奇曲率非负。进一步假设在某个奇异点x，局部群Γ_x 作用不可约，即Γ_x 的作用不存在非平凡不变子空间。证明：$b_1(M) = 0$。

下面是一些题目中出现的概念。一个n 维轨形M 是一个拓扑空间，在每个点x，存在邻域U_x 和一个同胚$\phi_x : U_x \to B/\Gamma_x$ 满足$\phi_x(x) = 0$，其中B 是n 为欧氏空间$\mathbb{R}^n$ 中的单位球，Γ_x 是$O(n)$ 的离散子群，在$\mathbb{R}^n$ 上具有线性、有效作用。组合(U_x, Γ_x, ϕ_x) 称为轨形的一个坐标卡。一个点x 称为正则点，如果Γ_x 是平凡的；否则，x 称为一个奇点。具有有限奇点$\Sigma = \{p_i\}_{i\in I}$ 的黎曼轨形M 是一个轨形，满足如下条件：(a) $M\backslash\Sigma$ 是一个光滑流形，赋予一个（非完备）黎曼度量g，(b) 在每个点p_i 存在坐标卡(U_i, Γ_i, ϕ_i) 使得(b1) ϕ_i 在$U_i\backslash\{p_i\}$ 上光滑，并且(b2) 在$U_i\backslash\{p_i\}$ 上，度量$g = \phi_i^* \tilde{g}_i$，其中$\tilde{g}_i$ 是B 上某个Γ_i 不变的光滑度量。M 称为完备的黎曼轨形，如果诱导的度量空间是完备的。

答案2

(1) 反设存在两个奇点P, Q. 令γ为连接他们最短路径. 则任意$x \in M$ 到γ 的最短距离在某点$y \in \gamma$取到. 若y 不为γ的端点，则从x 到y 的最短测地线与γ垂直。否则该测地线与γ的夹角不大于$\pi/2$ (奇点的性质). 由Heintze-Karcher 体积比较定理,

$$vol(B(r,\gamma)) \leq L(\gamma) \cdot \omega_{n-2} \cdot r^{n-1} \tag{3}$$

其中ω_{n-2} 为$(n-2)$-维单位球的体积. 从而无穷远处体积增长为0，矛盾。

(2) 任何$b_1(M)$ 中元素存在调和形式代表元ω. 由Bochner 公式,

$$\triangle|\omega|^2 = 2|\nabla\omega|^2 + 2Ric(\omega^\sharp, \omega^\sharp)$$

其中$\omega^\sharp$ 为ω的对偶向量场. 由极大值原理，ω 是平行的.

因此若$b_1 \neq 0$, 则存在非平凡的平行一形式ω. 在奇点x附近的坐标卡上, 余向量场$\omega(x)$ 在Γ_x作用下不变, 从而生成一个Γ_x的不变自空间，矛盾.

分析与方程

问题1

给定$\mathbb{R}^d$中任意n个点$P_1, P_2, \ldots, P_n$。

(1) 在直线$\mathbb{R}$上证明

$$\sum_{i<j}(-1)^{i+j}|P_i - P_j| \leq 0.$$

(2) 证明上述结论对任意维数$d \geq 1$都成立，这里$|P_i - P_j|$是欧式空间两点P_i与P_j之间的距离。

答案1

对于一维的情形，我们用归纳法证明。显然结论对$n \leq 2$成立。对$k + l$个点的情形，不妨假设$0 \leq x_1 \leq x_2 \leq \ldots \leq x_k$表示的是下标为奇数的点，$0 \leq y_1 \leq y_2 \leq \ldots \leq y_l$代表下标为偶数的点。显然$l = k$或者$l = k - 1$。由归纳假设($k + l - 2$的情形)只需要证明

$$\sum_{i<k}|x_k - x_i| + \sum_{j<l}|y_l - y_j| - \sum_{j<l}|x_k - y_j| - \sum_{i<k}|y_l - x_i| - |x_k - y_l| \leq 0.$$

由于$|l - k| \leq 1$，由大小关系可知

$$\begin{aligned}
&\sum_{j<l}|x_k - y_j| + \sum_{i<k}|y_l - x_i| + |x_k - y_l| \\
&\geq (l-1)x_k - \sum_{j<l} y_j + (k-1)y_l - \sum_{i<k} x_i + |x_k - y_l| \\
&= \sum_{i<k}|x_k - x_i| + \sum_{j<l}|y_l - y_j| + (l-k)(x_k - y_l) + |x_k - y_l| \\
&\geq \sum_{i<k}|x_k - x_i| + \sum_{j<l}|y_l - y_j|.
\end{aligned}$$

对于高维情形，利用投影化归为一维情形。记$\omega \in \mathbb{S}^{d-1}$为$\mathbb{R}^d$中的单位球面。记$P_i^\omega$是$P_i$到$O\omega$（通过原点与$\mathbb{S}^{d-1}$中的一点$\omega$的直线）的投影。对某个常数$C$ (只依赖于维数d)显然有

$$\int_{\mathbb{S}^{d-1}} |P_i^\omega - P_j^\omega| d\omega = C|P_i - P_j|.$$

实际上由球对称性不妨假设P_i, P_j落在x_1轴上。再由平移对称性，可假设P_i为原点。最后由伸缩不变性，可不妨假设P_j落在单位球面$\mathbb{S}^{d-1}$上。这表明上式中的C只依赖于维数d，与P_i, P_j的选取无关。这样

$$\sum_{i<j}(-1)^{i+j}|P_i - P_j| = C^{-1}\sum_{i<j}\int_{\mathbb{S}^{d-1}} |P_i^\omega - P_j^\omega| d\omega \leq 0.$$

问题2

对于复平面$\mathbb{C}$上的全纯函数$f(z)$，归纳定义$f^{(n+1)}(z) = f(f^{(n)}(z))$, $f^{(1)}(z) = f(z)$。是否存在$\mathbb{C}$上的全纯函数$f(z)$使得对任意$z \in \mathbb{C}$都有$f^{(2023)}(z) = e^{2023z}$？证明你的结论。

答案2

不存在。否则假设f是$\mathbb{C}$上的全纯函数满足$f^{(2023)}(z) = e^{2023z}$，特别的有像集$f(\mathbb{C})$为$C^* = \mathbb{C}\backslash\{0\}$。取$\mathbb{C}$上整函数$g(z)$使得$f(z) = e^{g(z)}$。这样可知存在整数$k$使得

$$g(f^{(2022)}(z)) = 2023z + 2k\pi i, \quad \forall z \in \mathbb{C}.$$

用$f(z)$代替z，可知$g(e^{2023z}) = 2023f(z) + 2k\pi i$。特别的

$$2023f(\mathbb{C}) + 2k\pi i = \mathbb{C}\backslash\{2k\pi i\} = g(e^{2023\mathbb{C}}) = \mathbb{C}\backslash\{g(0)\}.$$

这样可知$g(0) = 2k\pi i$，继而可知

$$g(f^{(2022)}(0)) = 2k\pi i = g(0)$$

由于f是单射，可要求g是单射，从而有$f^{(2022)}(0) = 0$，与$f(\mathbb{C}) = C^*$矛盾。

应用与计算数学

问题1

为了使大语言模型(LLM) 与人类偏好保持一致，我们需要对其进行微调。给定一个特定的prompt，让LLM生成n个回答。人类标注者将对这n 个回答进行从最好到最差的排序。用π 表示排序（比如，第$\pi(1)$ 个回答认为是最好的）。假设我们有一个在区间$[-1,1]$ 上强凹（等价地，$-G$是强凸）的且单调递增的函数G。我们的目标是训练一个奖励模型，为第i 个回答分配一个$0 \leq r_i \leq 1$ 的评分。理想情况下，奖励r_i 应该是以下优化问题的解：

$$\max_{0\leq r_1,\ldots,r_n\leq 1} \sum_{i<j} G(r_{\pi(i)} - r_{\pi(j)}).$$

(1) 解释为什么

$$L(r_1,\ldots,r_n) := \sum_{i<j} G(r_{\pi(i)} - r_{\pi(j)})$$

是凹的但不是强凹的（如果需要，可以假设G 具有足够的光滑性）。尽管如此，请证明上述优化问题仍具有唯一解$(r_1^\star,\ldots,r_n^\star) \in [0,1]^n$。

(2) 证明解必须满足

$$1 = r^\star_{\pi(1)} \geq r^\star_{\pi(2)} \geq \cdots \geq r^\star_{\pi(n)} = 0.$$

以及对于所有的$i = 1,\ldots,n$，证明$r^\star_{\pi(i)} + r^\star_{\pi(n+1-i)} = 1$。

(3) 现在考虑问题的极限版本。假设当$n \to \infty$ 时，$r_1^\star,\ldots,r_n^\star$ 的经验分布收敛于区间$[0,1]$上的概率测度μ。问题可以重新表述为

$$\sup_\mu \mathbb{E}_{X,X' \overset{iid}{\sim} \mu} G(|X - X'|),$$

其中X, X' 是从μ 独立且同分布抽取的。如果$\mathbb{E}_{X,X' \overset{iid}{\sim} \mu} G(|X - X'|)$在一个概率测度$\mu^\star$上取得最大值，试证明$\mathbb{E}_{X\sim\mu^\star} G(|X - c|)$ 不依赖于$c \in [0,1]$。

答案1

不失一般性，假设$\pi(i) = i$。令H为L的Hessian矩阵，则当$i \neq j$时，有

$$H_{ij} = -G''(r_{\min(i,j)} - r_{\max(i,j)})$$

当$i = j$时，有

$$H_{ii} = -\sum_{j\neq i} H(ij).$$

因此，H是对角占优矩阵，从而H是半正定的。唯一特征值为0的特征向量是全1向量，这表明唯一的零空间方向是平移。然而，由于G是增函数，有$r_1^\star = 1$且$r_n^\star = 0$，因此解是唯一的，(1)得证。

由于G的单调性，显然最优解必须具有最大间隔$r^\star_{\pi(1)} - r^\star_{\pi(n)}$，因此$r^\star_{\pi(1)} = 1$且$r^\star_{\pi(n)} = 0$。此外，如果$r^\star$是一个解，那么$1 - r^\star$也必须是解（反转排序）。这证明了(2)。

利用变分法。对于$\mathcal{B}([0,1])$上的两个概率测度μ和ν，有$(\mu-\nu)([0,1])=0$。假设μ满足$\mathbb{E}_{X\sim\mu}G(|X-c|)=K$不依赖于$c\in[0,1]$。注意到

$$\langle \nu-\mu,\mu\rangle=\int_{[0,1]}\left(\int_{[0,1]}|x-y|^\beta\mu(dx)\right)(\nu-\mu)(dy)=\int_{[0,1]}K(\nu-\mu)(dy)=0.$$

接下来，有$\langle \nu-\mu,\nu-\mu\rangle\le 0$。因此，

$$\langle \nu,\nu\rangle=\langle \mu,\mu\rangle+2\langle \nu-\mu,\mu\rangle+\langle \nu-\mu,\nu-\mu\rangle\le\langle \mu,\mu\rangle.$$

这证明了(3)。但是请注意，如果学生没有考虑μ的支撑严格小于$[0,1]$的情况，应该扣分。

问题2

考虑一个线性的Fokker-Planck 方程

$$\partial_t f=\Delta f+\nabla\cdot[f\nabla V(x)],\quad t>0,\ x\in\Omega\subset\mathbb{R}^d. \tag{4}$$

这里$f=f(t,x)\ge 0$是一个依赖于时间t和位置x的概率密度函数, $V(x)$是一个给定的函数在区域Ω里有下界。注意到Fokker-Planck 方程(4) 可以视为一个输运扩散方程。为数值求解(4)，我们引进函数$M(x)=\exp(-V(x))$，把它写成一个等价形式

$$\partial_t f=\nabla\cdot\left(M\nabla\left(\frac{f}{M}\right)\right). \tag{5}$$

这样的话，这个方程可以看成一个变系数的扩散方程，然后自然地用有限差分法来离散求解。具体来说，我们假定$d=1$, $\Omega=[a,b]$, 然后把$[a,b]$ 等分成N 个长度是$\Delta x=(a-b)/N$的小区间，区间中点是$x_i=a+(i-1/2)\Delta x$, $i=1,\cdots,N$，区间端点是$x_{i+1/2}=a+i\Delta x$, $i=0,\cdots,N$。对方程(5)在空间上使用中心差分，时间上使用向前Euler法就得到格式

$$\frac{f_i^{n+1}-f_i^n}{\Delta t}=\frac{F_{i+1/2}^n-F_{i-1/2}^n}{\Delta x},\quad F_{i+1/2}^n:=\frac{1}{\Delta x}M_{i+1/2}\left(\frac{f_{i+1}^n}{M_{i+1}}-\frac{f_i^n}{M_i}\right), \tag{6}$$

这里f_i^n是f在x_i 处和时间$t_n=n\Delta t$时对f的数值逼近，M_i 和$M_{i+1/2}$ 是M 分别在x_i 和$x_{i+1/2}$处的取值。在边界a和b处我们取$F_{1/2}^n=F_{N+1/2}^n=0$。

a) 对于数值格式(6)，证明能量泛函$E(f):=\int_\Omega f\ln\left(\frac{f}{M}\right)\mathrm{d}x$ 在离散意义下是单调下降的，即，证明

$$E_\Delta(f_i^{n+1})\le E_\Delta(f_i^n),\quad 这里\ E_\Delta(f_i^n):=\sum_{i=1}^N f_i^n\ln\left(\frac{f_i^n}{M_i}\right).$$

陈述清楚需要的条件。这个数值解会最终收敛到一个稳态解吗？是的话，写出稳态解的形式；不是的话，说明原因。

b) 我们还可以使用向后Euler法离散上面的方程，得到格式

$$\frac{f_i^{n+1}-f_i^n}{\Delta t}=\frac{F_{i+1/2}^{n+1}-F_{i-1/2}^{n+1}}{\Delta x}.$$

对这个格式，重复a)里的所有问题。

答案2

显式的数值格式可以被等价地写成

$$\frac{f_i^{n+1}}{M_i} = \left(1 - \frac{\Delta t}{\Delta x^2}\frac{M_{i+1/2}}{M_i} - \frac{\Delta t}{\Delta x^2}\frac{M_{i-1/2}}{M_i}\right)\frac{f_i^n}{M_i} + \frac{\Delta t}{\Delta x^2}\frac{M_{i+1/2}}{M_i}\frac{f_{i+1}^n}{M_{i+1}} + \frac{\Delta t}{\Delta x^2}\frac{M_{i-1/2}}{M_i}\frac{f_{i-1}^n}{M_{i-1}}. \quad (7)$$

定义$g_i^n = f_i^n/M_i$, $\mathbf{g}^n = (g_1^n, \dots, g_N^n)^T$, 则(7) 可以写成

$$\mathbf{g}^{n+1} = A\mathbf{g}^n, \quad (8)$$

其中A 是一个三对角矩阵满足

$$a_{ii} = 1 - \frac{\Delta t}{\Delta x^2}\frac{M_{i+1/2}}{M_i} - \frac{\Delta t}{\Delta x^2}\frac{M_{i-1/2}}{M_i}, \quad a_{i,i-1} = \frac{\Delta t}{\Delta x^2}\frac{M_{i-1/2}}{M_i}, \quad a_{i,i+1} = \frac{\Delta t}{\Delta x^2}\frac{M_{i+1/2}}{M_i}. \quad (9)$$

对矩阵A, 可以验证如下性质:

1. 通过选取$\Delta t \le \Delta x^2 \min_i \left\{\frac{M_i}{M_{i+1/2} + M_{i-1/2}}\right\}$, 可使$A \ge 0$, 换言之, A 中每一项都是非负的;
2. $A\mathbf{1} = \mathbf{1}$ ($\mathbf{1}$ 是全为1 的列向量), i.e., $\sum_j a_{ij} = 1$;
3. $A^T\mathbf{M} = \mathbf{M}$, 其中$\mathbf{M} = (M_1, \dots, M_N)^T$, i.e., $\sum_j M_j a_{ji} = M_i$.

由性质1 可知数值格式是保持正性的, 从而可以考虑熵:

$$\begin{aligned} &f_i^n \ge 0 \text{ 对所有} i \Longrightarrow g_i^n = f_i^n/M_i \ge 0 \text{ 对所有} i \\ \Longrightarrow &g_i^{n+1} = \sum_j a_{ij} g_j^n \ge 0 \text{ 对所有} i \Longrightarrow f_i^{n+1} = M_i g_i^{n+1} \ge 0 \text{ 对所有} i. \end{aligned} \quad (10)$$

结合性质1 与2 可知$g_i^{n+1} = \sum_j a_{ij} g_j^n$ 是一个凸组合从而可以应用Jensen 不等式(见下面).

由性质3 可知数值格式是质量守恒的:

$$\sum_i f_i^{n+1} = \sum_i M_i g_i^{n+1} = \sum_i M_i \left(\sum_j a_{ij} g_j^n\right) = \sum_j \left(\sum_i M_i a_{ij}\right) g_j^n = \sum_j M_j g_j^n = \sum_j f_j^n. \quad (11)$$

对离散熵, 有

$$\begin{aligned} E_\Delta(f_i^{n+1}) &= \sum_i f_i^{n+1} \ln\left(\frac{f_i^{n+1}}{M_i}\right) = \sum_i M_i g_i^{n+1} \ln g_i^{n+1} \le \sum_i M_i \left(\sum_j a_{ij} g_j^n \ln g_j^n\right) \\ &= \sum_j \left(\sum_i M_i a_{ij}\right) g_j^n \log g_j^n = \sum_j M_j g_j^n \ln g_j^n = \sum_j f_j^n \ln\left(\frac{f_j^n}{M_j}\right) = E_\Delta(f_i^n), \end{aligned} \quad (12)$$

其中我们用了Jensen 不等式来得到不等式估计. 从而数值格式也是熵减的.

进一步, $E_\Delta(f_i^n)$ 有下界, 从而序列$\{E_\Delta(f_i^n)\}$ 的极限存在. 对(12) 两边取极限可得$g_1^\infty = \cdots = g_N^\infty$. 再由质量守恒以及$M$ 的形式知, 稳定态有以下形式:

$$f_i^\infty = C\exp(-V(x_i)), \quad \text{其中} C = \frac{\sum_i f_i^0}{\sum_i \exp(-V(x_i))} \text{ 以及} f_i^0 \text{ 是初始条件.} \tag{13}$$

b) 显式的数值格式可以被等价地写成

$$\left(1 + \frac{\Delta t}{\Delta x^2}\frac{M_{i+1/2}}{M_i} + \frac{\Delta t}{\Delta x^2}\frac{M_{i-1/2}}{M_i}\right)\frac{f_i^{n+1}}{M_i} - \frac{\Delta t}{\Delta x^2}\frac{M_{i+1/2}}{M_i}\frac{f_{i+1}^{n+1}}{M_{i+1}} - \frac{\Delta t}{\Delta x^2}\frac{M_{i-1/2}}{M_i}\frac{f_{i-1}^{n+1}}{M_{i-1}} = \frac{f_i^n}{M_i}. \tag{14}$$

定义$g_i^n = f_i^n/M_i$, $\mathbf{g}^n = (g_1^n, \ldots, g_N^n)^T$, 则(14) 可以被写成

$$B\mathbf{g}^{n+1} = \mathbf{g}^n, \tag{15}$$

其中B 是一个三对角矩阵满足

$$b_{ii} = 1 + \frac{\Delta t}{\Delta x^2}\frac{M_{i+1/2}}{M_i} + \frac{\Delta t}{\Delta x^2}\frac{M_{i-1/2}}{M_i}, \quad b_{i,i-1} = -\frac{\Delta t}{\Delta x^2}\frac{M_{i-1/2}}{M_i}, \quad b_{i,i+1} = -\frac{\Delta t}{\Delta x^2}\frac{M_{i+1/2}}{M_i}. \tag{16}$$

对矩阵B, 可以验证如下性质:

1. B 是一个M 矩阵, 则$B^{-1} \geq 0$, i.e., B^{-1} 的每一项都是非负的;
2. $B\mathbf{1} = \mathbf{1}$, 则$B^{-1}\mathbf{1} = \mathbf{1}$;
3. $B^T\mathbf{M} = \mathbf{M}$, 则$(B^{-1})^T\mathbf{M} = \mathbf{M}$.

将数值格式(15) 视为

$$\mathbf{g}^{n+1} = B^{-1}\mathbf{g}^n. \tag{17}$$

对矩阵B^{-1}, 在显式的数值格式中, 它恰好满足与矩阵A 同样的性质. 从而, 正性, 质量守恒, 熵耗散, 以及稳定态都可以被容易地得到.

组合与概率

问题1

在一场舞会开始时，舞池中有20个女孩和22个男孩，舞池外还有无穷多的等待入场的女孩和男孩。舞池的游戏规则如下：每一轮机会均等地从舞池中选出一个幸运儿。如果被选出来的是名女孩，她会邀请一位舞池中的男孩跳一支双人舞，曲子结束以后双双离开舞池；如果被选出来的是名男孩，那么他会从舞池外等待的人中邀请一位女孩和一位男孩入场跳一支三人舞，舞曲结束后三人均留在舞池内。如果舞池里只剩两个男孩则舞会结束。

(1) 这场舞会永远也不会结束的概率是多少？

(2) 舞会的组织者决定颠倒舞会的规则：在每一轮中，如果被选出来的是名女孩，她会邀请会从舞池外等待的人中邀请一位女孩和一位男孩入场跳一支三人舞，舞曲结束后三人均留在舞池内；如果被选出来的是名男孩，那么他会邀请一位舞池中的女孩跳一支双人舞，曲子结束以后双双离场。舞会仍然在舞池里只剩两个男孩时结束。在新的规则下，该舞会平均会进行多少轮才会终止？

答案1

(1) 更一般地，我们考虑一场有n个女孩和$(n+2)$个男孩的舞会。我们将k轮舞曲后舞池内的女孩数记作X_k。令e_j为$X_0=j$时$(X_n)_{n\geq 0}$在有限步后碰到0的概率。由定义

$$e_j = p_j e_{j+1} + q_j e_{j-1},\ j \geq 1,\ \text{其中} p_j = \tfrac{j+2}{2(j+1)},\ q_j = \tfrac{j}{2(j+1)}\ \text{且} e_0 = 1\text{。}$$

根据马尔科夫链首达时的经典结果，e_j应该为该线性方程组最小的非负解。我们递归地给出e_j的表达式：

$$e_j = 1 - (1-e_1)\Big[1 + \frac{q_1}{p_1} + \cdots + \frac{q_1 q_2 \cdots q_{j-1}}{p_1 p_2 \cdots p_{j-1}}\Big].$$

容易发现，上式右侧方括号中的项求和后为$2-\frac{2}{j+1}$。因此，使得e_j的解均非负的解是$e_j = 1/(j+1)$。由此可知舞会永不结束的概率是$20/21$。

(2) 令d_j为“从有j个女孩和$(j+2)$个男孩的开场到$(j-1)$个女孩所需要的轮数”的期望。类似地，我们有如下递推式：

$$d_j = 1 + q_j[d_{j+1} + d_j].$$

令$f_j = d_j + (2j+1)$, 以上方程组可以简化为

$$(j+2)f_j = j f_{j+1}, \quad \text{于是}\ f_j = \frac{j(j+1)}{2} f_1.$$

整个舞会所需要的场数的期望则可记为：

$$s_n := \sum_{i=1}^{n} d_j = n(n+2) + f_1 \sum_{j=1}^{n} \frac{j(j+1)}{2}.$$

我们再次利用经典的马氏链首达时的结果：$(s_n)_{n\geq 0}$是以上线性方程组最小的非负解。显然最小解在$f_1=0$时取到。因此$s_{20}=420$。

问题2

给定正整数$k \geq 2$和，对充分大的正整数m,定义$\mathcal{F}_m$为包含所有恰好有m个1的0/1矩阵(不一定为正方阵)的集合. 定义$f(m)$为最大的正整数L,使得对$\mathcal{F}_m$中任意矩阵A, 都能找到一个同样大小的0/1矩阵B满足: (1) B至少有L个1; (2) B的每个元素都不大于A对应位置的元素; (3) B不包含任意k乘k的全1子矩阵. 证明等式

$$\lim_{m\to\infty} \frac{\ln f(m)}{\ln m} = \frac{k}{k+1}.$$

答案2

这个问题等价于以下图论问题: 对充分大的m, 找到最大的$f(m)$, 使得每个m条边的二分图G都包含一个子图H, 使得H至少有$f(m)$条边, 并且H不包含任何完全二分图$K_{k,k}$作为其子图. 我们将证明$f(m) = \Theta(m^{\frac{k}{k+1}})$. 立得极限

$$\lim_{m\to\infty} \frac{\ln f(m)}{\ln m} = \frac{k}{k+1}.$$

我们先证明下界$f(m) = \Omega(m^{\frac{k}{k+1}})$, 我们独立随机地以概率$p \in (0,1)$保留$G$里面的每条边, 称所得到的随机子图为$G'$. 定义两个随机变量: $X = e(G')$为G'的边数, Y 为G'里面$K_{k,k}$子图的个数. 由于$K_{k,k}$的顶点可被k条边的完美匹配覆盖，所以图G里面$K_{k,k}$子图的个数不超过m^k. 利用期望的线性性, 我们可以得到

$$\mathbb{E}[X] = pm,$$
$$\mathbb{E}[Y] \leq p^{k^2} m^k.$$

定义一个新的随机变量$Z = X - Y$. 不难观察到只要在G'中的Y个$K_{k,k}$子图每个里面去掉至多一条边, 我们就可以得到一个不包含$K_{k,k}$子图的新的图G'', 使得G''至少包含$X - Y$条边. 那么由期望的定义，一定存在一个这样的G''，使得其边数至少是$\mathbb{E}[Z] \geq pm - p^{k^2} m^k$. 我们选取$p = \frac{1}{k^{2/(k^2-1)} m^{1/(k+1)}}$ 使得这个下界尽可能大. 计算可得

$$e(G'') \geq \mathbb{E}[Z] = \Omega(m^{\frac{k}{k+1}}).$$

接下来我们证明上界$f(m) = O(m^{\frac{k}{k+1}})$，换句话说我们需要构造$m$条边的二分图$G$，使得其最大的不包含$K_{k,k}$的子图的边数至多为$O(m^{\frac{k}{k+1}})$. 对充分大的$m$, 令$G$为顶点集$A \cup B$上的完全二分图, 并满足$|A||B| = m$. 假定$H$是$G$的一个不包含$K_{k,k}$的子图. 那么$H$里面一个顶点在$A$里、剩下与其相邻的$k$个顶点在$B$里的$K_{1,k}$的数目等于(假定$d_v$是$H$里面顶点$v$的度数)

$$\sum_{v\in A} \binom{d_v}{k} \geq |A| \binom{m/|A|}{k}.$$

这里我们利用了$\binom{x}{k}$是一个凸函数. 对任意给定的B里面每个k个顶点的子集$T \subset B$, 令d_T为A里和所有T里的顶点在图H里相邻的顶点的个数. 由算两次可得

$$\sum_{T:T\subset B, |T|=k} d_T = \sum_{v\in A} \binom{d_v}{k} \geq |A| \binom{e(H)/|A|}{k},$$

另一方面, H里面$K_{k,k}$的数目等于

$$\sum_{T:T\subset B,|T|=k}\binom{d_T}{k}.$$

由于H不包含$K_{k,k}$, 我们马上得到对所有T, $d_T\le k-1$, 从而由上述不等式得到

$$|A|\binom{e(H)/|A|}{k}\le\binom{|B|}{k}(k-1),$$

化简不等式有

$$e(H)=O(|B||A|^{(k-1)/k}).$$

取$|A|=m^{k/(k+1)}$及$|B|=m^{1/(k+1)}$, 我们得到H边数的一个上界$e(H)=O(m^{k/(k+1)})$.

附注: 这个结果见于Conlon, Fox and Sudakov 的以下文章[Short proofs of some extremal results II, *J. Combin. Theory Ser. B* **121** (2016), 173–196.].