

不动点定理笔记

徐晨凯

2025 年 12 月 3 日

巴拿赫不动点定理

一、压缩映射

定义 1 (压缩映射). 设 (X, d) 是度量空间, $T : X \rightarrow X$ 是映射。如果存在常数 $k \in [0, 1)$, 使得对任意 $x, y \in X$, 都有

$$d(T(x), T(y)) \leq k \cdot d(x, y),$$

则称 T 是 X 上的压缩映射 (或收缩映射), 常数 k 称为压缩系数。

直观理解: 压缩映射将空间中任意两点之间的距离至少缩小为原来的 k 倍 (其中 $k < 1$)。这种“压缩”性质保证了映射不会使点与点之间的距离扩大。

例 1 (完整实数空间上的压缩映射). 考虑 $X = \mathbb{R}$, 配备通常的度量 $d(x, y) = |x - y|$ 。定义 $T : \mathbb{R} \rightarrow \mathbb{R}$ 为

$$T(x) = \frac{1}{2}x + 3.$$

对任意 $x, y \in \mathbb{R}$, 有

$$d(T(x), T(y)) = \left| \frac{1}{2}x + 3 - \frac{1}{2}y - 3 \right| = \frac{1}{2}|x - y| = \frac{1}{2} \cdot d(x, y).$$

等于成立小于等于也成立, 根据定义 T 是压缩映射, 压缩系数为 $k = \frac{1}{2}$ 。

例 2 (实数区间上的压缩映射). 设 $X = [0, 1]$, 配备通常的度量 $d(x, y) = |x - y|$ 。定义 $T : [0, 1] \rightarrow [0, 1]$ 为

$$T(x) = \frac{x}{3} + \frac{1}{4}.$$

首先验证 T 确实将 $[0, 1]$ 映入自身: 对任意 $x \in [0, 1]$,

$$T(x) = \frac{x}{3} + \frac{1}{4} \in \left[\frac{1}{4}, \frac{1}{3} + \frac{1}{4} \right] = \left[\frac{1}{4}, \frac{7}{12} \right] \subset [0, 1].$$

其次, 对任意 $x, y \in [0, 1]$,

$$|T(x) - T(y)| = \left| \frac{x}{3} - \frac{y}{3} \right| = \frac{1}{3}|x - y|.$$

如上例知道 T 是压缩映射, 压缩系数为 $k = \frac{1}{3}$ 。

例 3 (非压缩映射的例子). 考虑 $T : \mathbb{R} \rightarrow \mathbb{R}$ 定义为 $T(x) = 2x$ 。对任意 $x \neq y$,

$$|T(x) - T(y)| = |2x - 2y| = 2|x - y|.$$

由于 $2 > 1$, 不存在 $k \in [0, 1)$ 使得压缩条件成立, 因此 T 不是压缩映射。

二、度量空间中的连续映射

定义 2 (度量空间中的连续映射). 设 (X, d_X) 和 (Y, d_Y) 是两个度量空间, $f : X \rightarrow Y$ 是映射, $x_0 \in X$ 。如果对任意 $\varepsilon > 0$, 存在 $\delta > 0$, 使得当 $d_X(x, x_0) < \delta$ 时, 有

$$d_Y(f(x), f(x_0)) < \varepsilon.$$

称 f 在 x_0 处连续。

如果 f 在 X 的每一点都连续, 则称 f 是。

命题 1 (的等价定义). 设 (X, d_X) 和 (Y, d_Y) 是度量空间, $f : X \rightarrow Y$ 是映射。:

- (1) f 是连续映射 (在每一点连续)。
- (2) (序列刻画) 对任意 $x_0 \in X$ 和任意收敛到 x_0 的序列 $\{x_n\}$, 都有 $f(x_n) \rightarrow f(x_0)$ 。
- (3) (开集刻画) 对 Y 中的任意开集 V , $f^{-1}(V)$ 是 X 中的开集。
- (4) (闭集刻画) 对 Y 中的任意闭集 F , $f^{-1}(F)$ 是 X 中的闭集。

证明框架: 我们证明 $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (1)$ 。

$(1) \Rightarrow (2)$: 设 f 在每一点连续, $x_0 \in X$, $\{x_n\}$ 是收敛到 x_0 的序列。对任意 $\varepsilon > 0$, 由 f 在 x_0 连续, 存在 $\delta > 0$ 使得当 $d_X(x, x_0) < \delta$ 时, $d_Y(f(x), f(x_0)) < \varepsilon$ 。由于 $x_n \rightarrow x_0$, 存在 N 使得当 $n > N$ 时, $d_X(x_n, x_0) < \delta$ 。因此当 $n > N$ 时, $d_Y(f(x_n), f(x_0)) < \varepsilon$, 即 $f(x_n) \rightarrow f(x_0)$ 。

$(2) \Rightarrow (3)$: 设 $V \subseteq Y$ 是开集。需要证明 $U = f^{-1}(V)$ 是开集。任取 $x_0 \in U$, 则 $f(x_0) \in V$ 。因为 V 是开集, 存在 $r > 0$ 使得 $B_Y(f(x_0), r) \subseteq V$ 。我们用反证法证明存在 $\delta > 0$ 使得 $B_X(x_0, \delta) \subseteq U$ 。

假设不存在这样的 δ 。则对每个 $n \in \mathbb{N}$, 存在 $x_n \in B_X(x_0, 1/n)$ 但 $x_n \notin U$ 。显然 $x_n \rightarrow x_0$ 。由条件 (2), $f(x_n) \rightarrow f(x_0)$ 。因此存在 N 使得当 $n > N$ 时, $d_Y(f(x_n), f(x_0)) < r$, 即 $f(x_n) \in B_Y(f(x_0), r) \subseteq V$ 。这意味着 $x_n \in f^{-1}(V) = U$, 与 $x_n \notin U$ 矛盾。

$(3) \Rightarrow (4)$: 设 $F \subseteq Y$ 是闭集。则 $V = Y \setminus F$ 是开集。由条件 (3), $f^{-1}(V)$ 是开集。注意到

$$f^{-1}(F) = f^{-1}(Y \setminus V) = X \setminus f^{-1}(V).$$

因为 $f^{-1}(V)$ 是开集, 所以 $f^{-1}(F)$ 是闭集。

(4) \Rightarrow (1): 设 $x_0 \in X$, 需要证明 f 在 x_0 连续。对任意 $\varepsilon > 0$, 令 $F = \{y \in Y : d_Y(y, f(x_0)) \geq \varepsilon\}$ 。集合 F 是闭球的补集的闭包, 是闭集。由条件 (4), $f^{-1}(F)$ 是闭集。

因为 $f(x_0) \notin F$, 所以 $x_0 \notin f^{-1}(F)$ 。因此 $x_0 \in X \setminus f^{-1}(F)$, 这是一个开集。存在 $\delta > 0$ 使得 $B_X(x_0, \delta) \subseteq X \setminus f^{-1}(F)$ 。这意味着对任意 $x \in B_X(x_0, \delta)$, 有 $f(x) \notin F$, 即 $d_Y(f(x), f(x_0)) < \varepsilon$ 。因此 f 在 x_0 连续。□

命题 2 (压缩映射是连续的). 压缩映射必定是连续映射。

证明: 设 $T : X \rightarrow X$ 是压缩映射, 压缩系数为 $k \in [0, 1)$ 。对任意 $x_0 \in X$ 和任意 $\varepsilon > 0$, 取 $\delta = \varepsilon$ 。当 $d(x, x_0) < \delta$ 时,

$$d(T(x), T(x_0)) \leq k \cdot d(x, x_0) < k \cdot \delta < \delta = \varepsilon.$$

因此 T 在 x_0 连续。由于 x_0 是任意的, 所以 T 是连续映射。□

三、巴拿赫不动点定理

定理 1 (巴拿赫不动点定理). 设 (X, d) 是完备的度量空间, $T : X \rightarrow X$ 是压缩映射。则: T 存在唯一的不动点, 即存在唯一的 $x^* \in X$ 使得 $T(x^*) = x^*$ 。

证明:

第一步: 构造 Cauchy 序列

任取初值 $x_0 \in X$, 定义迭代序列

$$x_{n+1} = T(x_n), \quad n = 0, 1, 2, \dots$$

我们首先估计相邻两项之间的距离。由压缩性质,

$$\begin{aligned} d(x_2, x_1) &= d(T(x_1), T(x_0)) \leq k \cdot d(x_1, x_0), \\ d(x_3, x_2) &= d(T(x_2), T(x_1)) \leq k \cdot d(x_2, x_1) \leq k^2 \cdot d(x_1, x_0). \end{aligned}$$

对任意 $n \geq 1$, 有

$$d(x_{n+1}, x_n) \leq k^n \cdot d(x_1, x_0).$$

现在证明 $\{x_n\}$ 是 Cauchy 序列。对任意 $m > n \geq 0$, 利用一直展开的三角不等式,

$$\begin{aligned} d(x_m, x_n) &\leq d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \cdots + d(x_{n+1}, x_n) \\ &\leq k^{m-1} d(x_1, x_0) + k^{m-2} d(x_1, x_0) + \cdots + k^n d(x_1, x_0) \\ &= d(x_1, x_0) (k^n + k^{n+1} + \cdots + k^{m-1}) \\ &= d(x_1, x_0) \cdot k^n (1 + k + k^2 + \cdots + k^{m-n-1}). \end{aligned}$$

由于 $k \in [0, 1)$, 几何级数 $1 + k + k^2 + \dots$ 收敛, 且

$$1 + k + k^2 + \dots + k^{m-n-1} < \frac{1}{1-k}.$$

因此

$$d(x_m, x_n) \leq \frac{k^n}{1-k} d(x_1, x_0).$$

对任意 $\varepsilon > 0$, 由于 $k^n \rightarrow 0$ (因为 $0 \leq k < 1$), 存在 N 使得当 $n \geq N$ 时,

$$\frac{k^n}{1-k} d(x_1, x_0) < \varepsilon.$$

因此当 $m > n \geq N$ 时, $d(x_m, x_n) < \varepsilon$, 即 $\{x_n\}$ 是 Cauchy 序列。

第二步: 证明序列收敛到不动点

因为 (X, d) 是完备的度量空间,Cauchy 序列 $\{x_n\}$ 收敛, 设 $x_n \rightarrow x^* \in X$ 。

下面证明 x^* 是 T 的不动点。由 T 的连续性 (压缩映射必连续),

$$T(x^*) = T\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} T(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = x^*.$$

因此 $T(x^*) = x^*$, 即 x^* 是 T 的不动点。

第三步: 证明不动点的唯一性

假设 y^* 也是 T 的不动点, 即 $T(y^*) = y^*$ 。由压缩性质,

$$d(x^*, y^*) = d(T(x^*), T(y^*)) \leq k \cdot d(x^*, y^*).$$

因为 $k \in [0, 1)$, 上式意味着 $(1 - k)d(x^*, y^*) \leq 0$ 。由于 $1 - k > 0$ 和 $d(x^*, y^*) \geq 0$, 必有 $d(x^*, y^*) = 0$, 即 $x^* = y^*$ 。因此不动点是唯一的。

第四步: 误差估计

在第一步中, 我们已经证明了对任意 $m > n$,

$$d(x_m, x_n) \leq \frac{k^n}{1-k} d(x_1, x_0).$$

令 $m \rightarrow \infty$, 由度量的连续性,

$$d(x^*, x_n) = \lim_{m \rightarrow \infty} d(x_m, x_n) \leq \frac{k^n}{1-k} d(x_1, x_0).$$

这给出了迭代序列到不动点的误差估计。□

直观理解: 巴拿赫不动点定理告诉我们, 在完备度量空间中, 压缩映射通过不断迭代, 会使任意初始点逐渐“聚拢”到唯一的不动点。压缩系数 $k < 1$ 保证了每次迭代都将距离缩小, 而完备性保证了这个“聚拢”的极限点确实存在于空间中。误差估计式表明收敛速度是指数级的, 且 k 越小, 收敛越快。

1 Brouwer 不动点定理

以下的内容都不是我原创的，我是看 B 站 [findingnothin](#) 和 [xan_derd](#) 两个很厉害的博主的基础上总结整理的，补充了一点我自己理解中有点糊涂地方的说明和一些直观的理解。

我们首先给出 Brouwer 不动点定理的表述。

定理 2 (Brouwer 不动点定理). 令 $D \subset \mathbb{R}^n$ 为非空、紧致且凸的集合。若映射 $f : D \rightarrow D$ 连续，则必存在一点 $x^* \in D$ ，使得

$$f(x^*) = x^*.$$

虽然“非空、紧致且凸”的表述已经很清晰，但在实际操作中仍显得过于抽象。在实数空间中，有一类特别好处理的紧致凸集——标准单纯形。它的定义如下：

定义 3 (标准单纯形). 对于正整数 n , \mathbb{R}^{n+1} 中的标准 n -单纯形定义为

$$\Delta^n = \left\{ x = (x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_i \geq 0, \sum_{i=1}^{n+1} x_i = 1 \right\}.$$

它是一个非空、紧致且凸的集合，并且是凸分析与代数拓扑中最常用的紧致凸集之一。

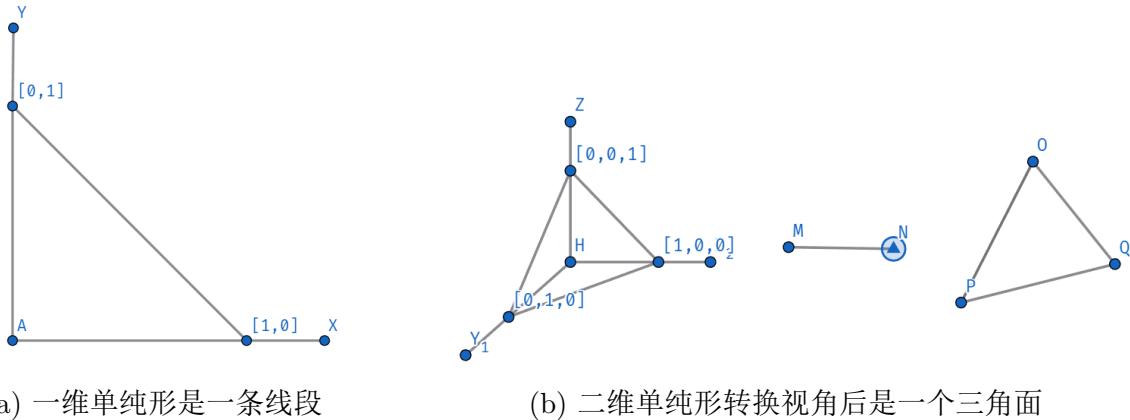


图 1: 单纯形的例子

现在我们的策略是：把一般情形下的紧致凸集，借助一个设计得当的映射“搬运”到标准单纯形上。直观地说，虽然不同的紧致凸集形状各异，但**标准单纯形**结构简单、定义清晰。如果我们能先证明：只要 Brouwer 定理在标准单纯形上成立，那么它在所有紧致凸集中也必然成立，那就等于把所有难题都缩减到一个最容易处理的模型空间中去。这样，“原先”在任意紧致凸集中找不动点”这一看起来很普遍的问题，就被转化成“只需要在标准单纯形上证明一次”这样一个更清爽、可操作性更强的任务了。

沿着这个思路，我们先来建立从一般紧致凸集到标准单纯形的约化。具体而言，我们要证明：

命题 3 (约化到标准单纯形). 如果 Brouwer 不动点定理在标准单纯形 Δ^n 上成立, 即对任意连续映射 $f : \Delta^n \rightarrow \Delta^n$ 都存在 $x^* \in \Delta^n$ 使得 $f(x^*) = x^*$, 那么对于任意非空、紧致且凸的集合 $D \subset \mathbb{R}^n$, 只要映射 $f : D \rightarrow D$ 连续, 也必定存在一点 $x^* \in D$ 使得 $f(x^*) = x^*$ 。

证明. 为了在 D 中构造不动点, 我们引入辅助映射族

$$F_\varepsilon(x) = \varepsilon x + (1 - \varepsilon)f(x), \quad \forall \varepsilon \in \mathbb{R}.$$

由于 D 的凸性以及 $x, f(x) \in D$, 对任意 $x \in D$ 都有 $F_\varepsilon(x) \in D$ 。同时, 由连续映射复合仍为连续的性质可知, $F_\varepsilon(x)$ 是连续的。

接下来我们在 D 中嵌入一个标准单纯形。回顾标准单纯形在 \mathbb{R}^{n+1} 中的定义:

$$\Delta^n = \left\{ (t_0, \dots, t_n) \mid \sum_{i=0}^n t_i = 1, t_i \geq 0 \ \forall i = 0, 1, \dots, n \right\}.$$

从 D 中取 $(n+1)$ 个仿射无关的点 $\{v_0, v_1, \dots, v_n\} \subset D$, 它们所张成的集合构成一个 n 维单纯形。定义映射

$$\mathcal{Y} : \Delta^n \rightarrow D, \quad \mathcal{Y}(z) = \sum_{i=0}^n t_i v_i, \quad \forall z = (t_0, \dots, t_n) \in \Delta^n.$$

注意到 $\mathcal{Y}(t_0, \dots, t_n) = \sum_{i=0}^n t_i v_i$ 是一个仿射映射, 在有限维空间中可以看成线性映射加上常向量, 因此是连续的。由于点集 $\{v_0, \dots, v_n\}$ 仿射无关, Δ^n 中的每一个凸组合都有唯一的系数表示, 所以 \mathcal{Y} 在其像上是一一对应的; 而它的像正好是由这些点构成的单纯形, 因此 \mathcal{Y} 也是满射。再结合 Δ^n 的紧致性 (连续双射把紧集映成紧集) 与欧氏空间的 Hausdorff 性 (紧子集就是闭子集), 可得连续双射的逆映射也是连续的, 于是 \mathcal{Y}^{-1} 也连续。

有了这个连续双射, 我们可以把 D 上的映射“拉回”到标准单纯形上。定义

$$\widehat{F}_\varepsilon : \Delta^n \rightarrow \Delta^n, \quad \widehat{F}_\varepsilon = \mathcal{Y}^{-1} \circ F_\varepsilon \circ \mathcal{Y}.$$

这里映射的顺序是: 把 Δ^n 上的点通过 \mathcal{Y} 映射到 D 上, 再通过 F_ε 映射到 D 上, 最后通过 \mathcal{Y}^{-1} 映射回 Δ^n 。由于连续映射的复合仍然是连续映射, 且 \mathcal{Y} 、 \mathcal{Y}^{-1} 和 F_ε 都是连续的, 因此 \widehat{F}_ε 是定义在 Δ^n 上的连续映射。

现在我们可以应用假设了。根据 Brouwer 不动点定理在 Δ^n 上成立 (这是我们的假定, 后面会证明这一点), 存在 $\widehat{x}_\varepsilon \in \Delta^n$ 使得

$$\widehat{F}_\varepsilon(\widehat{x}_\varepsilon) = \widehat{x}_\varepsilon.$$

将这个不动点映回 D , 令 $x_\varepsilon = \mathcal{Y}(\hat{x}_\varepsilon)$ 。由 $\hat{F}_\varepsilon = \mathcal{Y}^{-1} \circ F_\varepsilon \circ \mathcal{Y}$ 得到

$$\mathcal{Y}^{-1}(F_\varepsilon(x_\varepsilon)) = \hat{F}_\varepsilon(\hat{x}_\varepsilon) = \hat{x}_\varepsilon = \mathcal{Y}^{-1}(x_\varepsilon).$$

由于 \mathcal{Y}^{-1} 为单射, 因此

$$F_\varepsilon(x_\varepsilon) = x_\varepsilon.$$

将 F_ε 的定义代入, 由 $F_\varepsilon(x_\varepsilon) = x_\varepsilon$ 得

$$\varepsilon x_\varepsilon + (1 - \varepsilon)f(x_\varepsilon) = x_\varepsilon.$$

移项得

$$(1 - \varepsilon)f(x_\varepsilon) = (1 - \varepsilon)x_\varepsilon.$$

当 $\varepsilon \neq 1$ 时, 两边同时除以 $1 - \varepsilon$, 得到

$$f(x_\varepsilon) = x_\varepsilon.$$

这就是我们要找的不动点。(注: 当 $\varepsilon = 1$ 时, $F_1(x) = x$ 对任意 $x \in D$ 恒成立, 方程不能提供关于 f 的额外信息, 因此不会影响或改变我们对 f 不动点的推导。) \square

有了这个约化结果, 我们的任务就清晰了: 只需证明 Brouwer 不动点定理在[标准单纯形](#) Δ^n 上成立。为了实现这一目标, 我们需要引入 KKM 定理, 而要证明 KKM 引理, 我们又需要用到 Sperner 引理。为了更容易理解 Sperner 引理的作用, 我们先从一维的情况入手做一个直观理解。

定理 3 (一维 Sperner 引理). 设标准单纯形 $\Delta^1 = \{(t_0, t_1) \mid t_0 + t_1 = 1, t_0, t_1 \geq 0\}$ 被划分为若干小区间, 并在每个划分点上标记标签 0 或 1, 满足: 顶点 $(1, 0)$ 的标签为 0, 顶点 $(0, 1)$ 的标签为 1, 中间点的标签可以自由选择 0 或 1。则必然存在奇数个小区间, 其两个端点的标签不同, 即一端点为 0, 另一端点为 1。

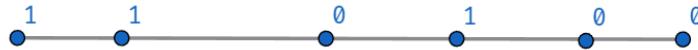


图 2: 其中一种可能

证明. 我们用两种方法来统计标签 1 的数量, 然后比较结果。

首先, 注意到 Δ^1 的右端点所包含的 1 只被计算一次, 而其他的 1 都被相邻的两个小区间共享, 因此它们被计算了两次。由此可知, 标签 1 的总数是奇数。

另一方面, 定义小区间中标签组合”01”和”10”的数量为 k , 标签组合”11”的数量为 l 。根据上述统计, 标签 1 的总数可以表示为 $k + 2l$ 。

由于两种方法统计的是同一个量，而第一种方法告诉我们总数为奇数，因此 $k + 2l$ 为奇数。注意到 $2l$ 是偶数，这推出 k 为奇数，从而至少存在一个标签组合为”01”或”10”的小区间。 \square

这个定理在直观上很好理解。如果把 0 理解成红色，1 理解成绿色，那么上述结果可以理解为：若两个端点染不同颜色，线段内部的顶点任意染这两种颜色，那么必然存在至少一对相邻顶点颜色不同。从左端（不妨为红）走到右端（对应为绿），中途一定会有一次”颜色切换”——比如从红点走到绿点，这两个相邻顶点就构成了”异色边”。哪怕中间有很多顶点，也至少会有一次切换（不可能从红直接变绿而不经过中间的颜色过渡点）。

这个证明特别优美，其中通过两种方法”计数”的思想在一维中就能看得很清楚。而这种思路在高维 Sperner 引理的证明中也会用到，这也是我们先介绍一维情况的主要原因。

顺着这个思路，我们来证明高维 Sperner 引理。在此之前，我们需要引入几个技术性定义。

定义 4 (面). 对任意集合 $T \subset S$ ，记 $\Delta(T)$ 为 $\Delta(S)$ 的 $|T| - 1$ 维面。

举个例子，设 $S = \{(0, 0, 1), (0, 1, 0), (1, 0, 0)\}$ 。由这三个点构成的单纯形是一个二维的三角形，如图一右侧所示；其中的每条边和每个顶点都属于这一结构。

如果我们只取其中两个点，例如 $(1, 0, 0)$ 和 $(0, 1, 0)$ ，那么所得的集合就不再形成完整的三角形，而是退化为连接这两点的一条线段，这与图一左侧的情况一致。该线段是一维结构，可以看作原二维单纯形的一个较低维“面”。

定义 5 (Carrier). 对于任意点 $x \in \Delta(S)$ ，它可以表示为凸组合

$$x = \sum_{i=1}^{m+1} \alpha_i a^i, \quad \alpha_i \geq 0, \quad \sum_{i=1}^{m+1} \alpha_i = 1.$$

定义 $C(x) = \{i \in \{1, \dots, m+1\} \mid \alpha_i > 0\}$ 为 x 的 *carrier*。

点 E 位于线段 AB 上，因此可以写成 A 与 B 的凸组合，即 $E = \lambda A + (1 - \lambda)B$ ，其中 $\lambda \in [0, 1]$ 。也就是说， E 的 carrier 仅由 $\{A, B\}$ 组成。相比之下，点 F 位于三角形 ABC 的内部，它可以表示为三个顶点的凸组合： $F = \alpha A + \beta B + \gamma C$ ，其中 $\alpha, \beta, \gamma > 0$ 且 $\alpha + \beta + \gamma = 1$ 。因此， F 的 carrier 是整个集合 $\{A, B, C\}$ 。

定义 6 (单纯形划分). 一个 m 维单纯形 $\Delta(S)$ 的单纯形划分 (*simplicial subdivision*) 是有限个 m 维小单形

$$\Upsilon = \{\Delta(T_1), \dots, \Delta(T_k)\}$$

的集合，满足 $\bigcup_{i=1}^k \Delta(T_i) = \Delta(S)$ ，并且对 $i \neq j$ ， $\Delta(T_i) \cap \Delta(T_j)$ 要么为空，要么是两者的公共面。记 $V(\Upsilon)$ 为所有小单形的顶点集合。

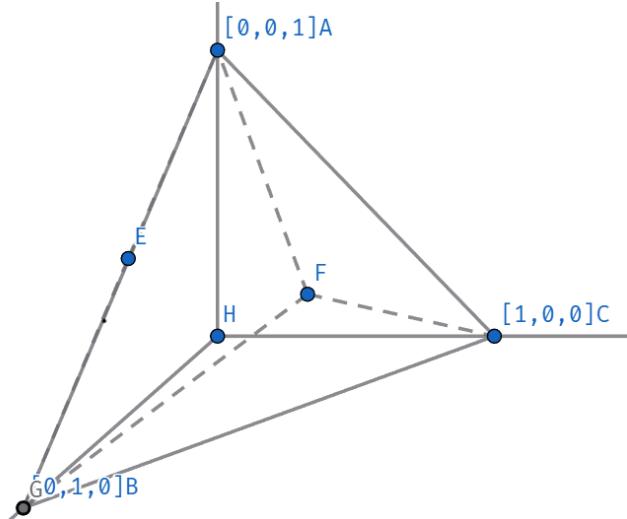


图 3: carrier 的示例

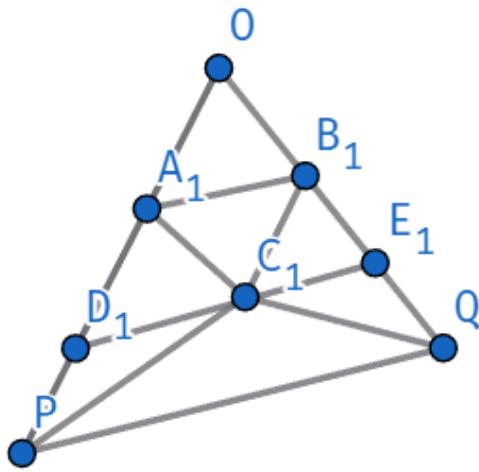


图 4: 单纯形分割的例子

这里的每一个小三角形，都是对大三角形的二维**单纯形划分**。把这些小三角形拼在一起，就组成了原来的大三角形。两个小三角形相交时，它们的公共部分只可能是三种情况：一个点（也就是 0 维的面）、一条线段（1 维的面），或者什么都没有。结合前面“面”的定义，我们可以看得很清楚：这些被划分出来的小单形之间的公共部分，其维度一定严格小于原来那个三角形的维度。 $V(\Upsilon)$ 就是上文中的所有字母的集合。

定义 7 (Sperner 标记). 设 $\Delta(S)$ 为 m 维单纯形， Υ 为其单纯形划分。对 Υ 的一个 Sperner 标记是映射

$$\lambda : V(\Upsilon) \rightarrow \{1, 2, \dots, m+1\},$$

并满足以下条件：

1. 若 v 为 $\Delta(S)$ 的一个顶点，则 $\lambda(v)$ 等于该顶点对应的编号；

2. 若 v 位于 $\Delta(S)$ 的某个子面 $\Delta(T)$ 上, 则 $\lambda(v)$ 必须属于 T ;
3. 若 v 在 $\Delta(S)$ 的内部, 则 $\lambda(v)$ 可以取任意值于 $\{1, 2, \dots, m+1\}$ 。

换言之, 每个顶点 $v \in V(\Upsilon)$ 的标签 $\lambda(v)$ 都必须属于其载体 $C(v)$ 。

其实理解起来并不困难。回头看我们讨论过的一维情形就能体会到, 这里的做法与当时完全一致: 顶点的标记仍然由它的 carrier 决定, 只不过在一维情况下 carrier 只有 0 和 1, 而现在扩展到了更高维的情况。

有了这些铺垫, 我们可以给出高维版本的 Sperner 引理。

定理 4 (Sperner 引理, 高维). 设 Δ^m 为 m 维单纯形, Υ 是其任意单纯剖分。若在 Υ 上给定一个满足 Sperner 条件的标记

$$\lambda : V(\Upsilon) \rightarrow \{1, 2, \dots, m+1\},$$

则 Υ 中至少存在一个完全标号的小单形(即其各顶点的标签恰好构成集合 $\{1, 2, \dots, m+1\}$)。更强地, 这样的完全标号小单形的数量一定是奇数。

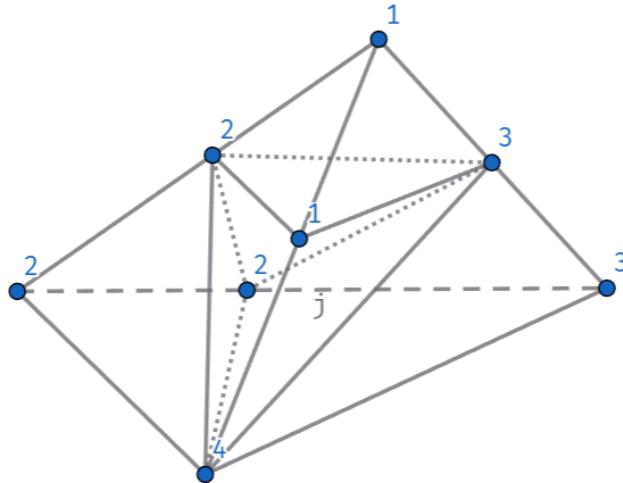


图 5: 看最中间的小单纯形

这是一个三维单纯形, 我随便赋的点, 你可以发现最中间的小三角体是被完全标号的。

证明. 我们用归纳法证明该引理。

当 $m = 0$ 时, 单纯形退化为一个点, 完全标号小单形就是他本身显然存在, 结论成立。

假设 Sperner 引理对 m 维单纯形成立, 下面证明对 $(m+1)$ 维单纯形也成立。设 $\Delta(S)$ 为 $(m+1)$ 维单纯形, 并考虑其单纯形划分 Υ 。定义集合

$$C = \{\text{顶点标签为 } \{1, 2, \dots, m+2\} \text{ 的小单形}\}, \quad A = \{\text{顶点标签为 } \{1, 2, \dots, m+1\} \text{ 的小单形}\},$$

$E = \{\text{所有位于单纯形} \Delta(S) \text{ 边界上, 且带有标签}\{1, 2, \dots, m+1\} \text{ 的} m \text{ 维面}\}.$

另外, 定义

$$R = C \cup A \cup E.$$

对于任意 $t, t' \in R$ 定义

$$D = \{(t, t') \mid t \neq t', \lambda(V(t \cap t')) = \{1, 2, \dots, m+1\}\}.$$

注意到如果 $(t, t') \in D$, 则 $(t', t) \in D$, 因此 $|D|$ 是偶数。

同时可以验证:

$$|\{t' \mid (t, t') \in D\}| = \begin{cases} 1, & t \in C, \\ 2, & t \in A, \\ 1, & t \in E. \end{cases}$$

其中:

- 若 $t \in C$, 由于完全标号的小单形恰好多出一个标签 $m+2$, 去掉该标签后只会形成唯一一个顶点标签为 $\{1, \dots, m+1\}$ 的面, 因此它只能找到一个邻居;
- 若 $t \in A$, 小单形维度未改变, 顶点标签集合恰好为 $\{1, \dots, m+1\}$, 其中必然至少有一个标签被重复使用, 因此可以找到两个共享该标签集合的相邻小单形;
- 若 $t \in E$, 若 $t \in E$, 该小单形位于边界, 它只能找到一个包含 1 到 m 的面的邻居。若尝试再构造第二个邻居, 例如在图五中 1-1-4 线段的 1-1 处插入一个标签 4, 理论上可以多出一个邻居, 但你可以验证, 这样的分割不再是一个合法的单纯形划分。这里的讨论非常不严谨, 但我无法想象更高维的情况了。这里需要区分: 我们讨论的不是完全标号的小单形数量, 而是一个特定属于 E 的 m 维面能够找到的邻居。

我们来看图五来理解这个过程。大三角锥是一个 $\Delta(4)$ 。最中间的那个小三角锥属于集合 C , 它能找到的“123 面”只有左侧那一面, 它找到的邻居是最上面的小三角形。

再看最上面的那个三角锥, 它可以提供两个“123 面”, 分别是底面和后面的面, 因此可以贡献两个不同的邻居。最后, 我们观察横切的“123 三角面”, 它看似贡献了两个面, 但这两个面实际上与之前 A 和 C 中的小单形重复使用了, 这并不是理论出错, 而是因为它不在边界上。

真正属于集合 E 的, 是边界上的大三角面和小三角面“123”, 它们各自只贡献一个邻居。

需要注意的是, D 是一个笛卡尔积集合, 调换元素顺序 (t, t') 后就对应不同的元素。于是全体 D 的大小可以表示为

$$|D| = |C| + 2|A| + |E|.$$

根据归纳假设, $|E|$ 是奇数, 而 $2|A|$ 是偶数, $|D|$ 是偶数, 由此推出 $|C|$ 必然是奇数。因此至少存在一个完全标号的小单形。

归纳完成, 高维 Sperner 引理成立。 \square

现在我们有了高维标准单纯形上的 Sperner 定理, 下一步是建立 KKM 引理。这个引理将为我们提供证明 Brouwer 不动点定理所需的关键工具。

定理 5 (Knaster–Kuratowski–Mazurkiewicz 引理). 设 Δ^{n-1} 是 $n-1$ 维标准单纯形, 其顶点为 e^1, \dots, e^n 。设 $F_1, \dots, F_n \subset \Delta^{n-1}$ 是闭集, 且满足以下条件:

$$\text{conv}\{e^i \mid i \in I\} \subset \bigcup_{i \in I} F_i, \quad \forall I \subset \{1, \dots, n\}, I \neq \emptyset,$$

即任意顶点子集生成的凸包都被对应指标的集合覆盖。

那么, 这些闭集的交集非空:

$$\bigcap_{i=1}^n F_i \neq \emptyset.$$

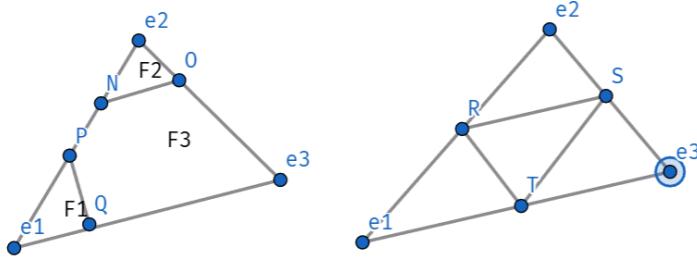


图 6: 左边不符合 KKM 要求, 右边符合

左边的 PN 不可以用 e_1 和 e_2 来表示, 不符合要求, 右边很完美。

证明. 对每个正整数 k , 取 Δ^{n-1} 的单纯形划分 Υ_k , 使得每个小单形 $T \in \Upsilon_k$ 的直径满足 $\sup_{x,y \in T} d(x,y) < 1/k$ 。记 $V(\Upsilon_k)$ 为划分中所有小单形的顶点集合。

对于每个顶点 $v \in V(\Upsilon_k)$, 由条件可知

$$v \in \text{conv}\{e^i \mid i \in C(v)\} \subset \bigcup_{i \in C(v)} F_i,$$

因此存在 $i \in C(v)$ 使得 $v \in F_i$ 。将这样的 i 作为顶点 v 的标签 $\lambda_k(v)$, 得到映射 $\lambda_k : V(\Upsilon_k) \rightarrow \{1, \dots, n\}$, 它满足 Sperner 标记条件。

由 Sperner 引理, Υ_k 中存在一个完全标号的小单形, 其顶点记为 $v^1(k), \dots, v^n(k)$, 满足 $\lambda_k(v^i(k)) = i$ 。根据标签的构造, 有 $v^i(k) \in F_i$ 对所有 i 成立。

由于 Δ^{n-1} 是紧致的, 序列 $\{v^i(k)\}_{k=1}^\infty$ 存在收敛子列 $\{v^i(k_m)\}_{m=1}^\infty$, 其极限记为 x_i^* 。当 $k \rightarrow \infty$ 时, $\sup_{x,y \in T} d(x,y) \rightarrow 0$, 因此小单形中的所有 $v^i(k_m)$ (对所有 i) 都收敛到

同一个极限点 x^* 。由于每个 F_i 是闭集，得到 $x^* \in F_i$ 对所有 i 成立。于是存在

$$x^* \in \bigcap_{i=1}^n F_i. \quad \square$$

终于，我们可以回到主线了。有了 KKM 引理作为工具，下一步是证明**标准单纯形**上的 Brouwer 不动点定理。

定理 6 (Brouwer 不动点定理, Δ^{n-1} 情况). 设 $f : \Delta^{n-1} \rightarrow \Delta^{n-1}$ 是**连续映射**，则存在 $x^* \in \Delta^{n-1}$ 满足

$$f(x^*) = x^*.$$

证明. 将映射写为坐标形式 $f(x) = (f_1(x), \dots, f_n(x))^T$ ，并对每个 $i = 1, \dots, n$ 定义闭集

$$F_i = \{x \in \Delta^{n-1} \mid x_i \geq f_i(x)\}.$$

我们的策略是：如果能证明 $\bigcap_{i=1}^n F_i \neq \emptyset$ ，那么交集中的任意点都是不动点。实际上，若存在 $x^* \in \bigcap_{i=1}^n F_i$ ，则对所有 i 有 $x_i^* \geq f_i(x^*)$ 。由于 $x^* \in \Delta^{n-1}$ ，有 $\sum_{i=1}^n x_i^* = 1$ ；又由于 $f(x^*) \in \Delta^{n-1}$ ，有 $\sum_{i=1}^n f_i(x^*) = 1$ 。由此可得

$$1 = \sum_{i=1}^n x_i^* \geq \sum_{i=1}^n f_i(x^*) = 1,$$

因此必须有 $x^* = f(x^*)$ ，也就是说， x^* 是不动点。

为了证明交集非空，我们要用 KKM 引理。首先需要验证 KKM 定理运用的前置条件。考虑任意 $I \subset \{1, \dots, n\}$ ，取 $x \in \text{conv}\{e^i \mid i \in I\}$ 。我们用反证法说明 $x \in \bigcup_{i \in I} F_i$ 。

若 $x \notin \bigcup_{i \in I} F_i$ ，则 $x_i < f_i(x)$ 对所有 $i \in I$ 成立。而由于 $x \in \text{conv}\{e^i \mid i \in I\}$ ，有 $x_i = 0$ 对所有 $i \notin I$ ，因此 $\sum_{i \in I} x_i = \sum_{i=1}^n x_i = 1$ 。另一方面， $\sum_{i \in I} f_i(x) \leq \sum_{i=1}^n f_i(x) = 1$ ，于是得到

$$1 = \sum_{i \in I} x_i < \sum_{i \in I} f_i(x) \leq 1,$$

矛盾。因此 $x \in \bigcup_{i \in I} F_i$ ，即满足 KKM 条件。

由 KKM 引理，闭集 F_1, \dots, F_n 的交集非空，记 $x^* \in \bigcap_{i=1}^n F_i$ 。根据前述分析， x^* 就是 f 的不动点。 \square

综合本节开头的约化命题，我们已经完整地证明了 Brouwer 不动点定理：从一般紧致凸集约化到**标准单纯形**，再利用 Sperner 引理建立 KKM 引理，最后在标准单纯形上构造不动点。

2 有限博弈中混合策略 Nash 均衡的存在性

考虑一个有限的正常型博弈 $(I, (A_i)_{i \in I}, (u_i)_{i \in I})$, 其中 I 是玩家集合, A_i 是玩家 i 的有限纯策略集合, u_i 是玩家 i 的收益函数。在实际博弈中, 玩家往往不会只选择某一个固定的纯策略, 而是会按照一定的概率分布在多个纯策略之间进行随机选择, 这种策略称为混合策略。

对于每个玩家 i , 我们将其混合策略集合记为

$$\Delta(A_i) = \left\{ \sigma_i \in [0, 1]^{|A_i|} : \sum_{a_i \in A_i} \sigma_i(a_i) = 1 \right\},$$

这里 $\sigma_i(a_i)$ 表示玩家 i 选择纯策略 a_i 的概率。这个集合实际上就是将纯策略按照概率混合所得到的概率分布空间。将所有玩家的混合策略组合在一起, 记为

$$\Sigma = \prod_{i=1}^n \Delta(A_i),$$

它是有限维欧氏空间中的一个紧致且凸的集合, 因为它是有限维单纯形的乘积。我们要证明的是:

定理 7 (Nash 均衡存在性定理). 在任何有限正常型博弈中, 混合策略 Nash 均衡必然存在。

在证明这个定理之前, 我们需要明确什么是 Nash 均衡。

定义 8 (Nash 均衡). 混合策略剖面 $\sigma^* = (\sigma_1^*, \dots, \sigma_n^*) \in \prod_{j=1}^n \Delta(A_j)$ 是博弈 $(I, (A_i)_{i \in I}, (u_i)_{i \in I})$ 的一个 Nash 均衡, 如果对所有玩家 $i \in I$, 都有

$$u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i, \sigma_{-i}^*)$$

对任意 $\sigma_i \in \Delta(A_i)$ 成立。

这个定义的含义是: 在 Nash 均衡中, 没有任何玩家能够通过单方面改变自己的策略来提高收益。需要注意的是, 上述定义要求对所有可能的混合策略 σ_i 都不能提高收益, 这看起来是一个很强的条件。然而, 由于期望收益的线性性质, 我们有如下关键观察:

命题 4. *Nash 均衡的等价刻画* 混合策略剖面 σ^* 是 Nash 均衡, 当且仅当对所有玩家 $i \in I$, 都有

$$u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(a_i, \sigma_{-i}^*)$$

对任意纯策略 $a_i \in A_i$ 成立。

证明. 必要性是显然的, 因为纯策略是混合策略的特例。

对于充分性，假设对所有纯策略 $a_i \in A_i$, 都有

$$u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(a_i, \sigma_{-i}^*).$$

对于任意混合策略 $\sigma_i \in \Delta(A_i)$, 由于 $\sigma_i(a_i) \geq 0$, 我们可以在上述不等式两边同时乘以 $\sigma_i(a_i)$, 得到

$$\sigma_i(a_i)u_i(\sigma_i^*, \sigma_{-i}^*) \geq \sigma_i(a_i)u_i(a_i, \sigma_{-i}^*).$$

对所有 $a_i \in A_i$ 求和, 得到

$$\sum_{a_i \in A_i} \sigma_i(a_i)u_i(\sigma_i^*, \sigma_{-i}^*) \geq \sum_{a_i \in A_i} \sigma_i(a_i)u_i(a_i, \sigma_{-i}^*).$$

注意到左边等于 $u_i(\sigma_i^*, \sigma_{-i}^*) \sum_{a_i \in A_i} \sigma_i(a_i) = u_i(\sigma_i^*, \sigma_{-i}^*)$, 而右边恰好是 $u_i(\sigma_i, \sigma_{-i}^*)$ 的定义, 因此

$$u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i, \sigma_{-i}^*).$$

这就证明了 σ^* 满足 Nash 均衡的定义。 \square

这个引理告诉我们: 要验证一个混合策略剖面是否为 Nash 均衡, 我们只需要检验它对所有纯策略的偏离是否都不能改进收益。这个等价刻画大大简化了问题, 因为我们只需要考虑有限多个纯策略, 而不需要考虑无限多个混合策略。

有了这个等价刻画, 我们的证明策略就清晰了: 构造一个从 Σ 到自身的连续映射, 然后利用 Brouwer 不动点定理找到不动点, 最后证明这个不动点满足上述引理中的条件。为了实现这一构造, 我们首先需要定义一些基本的量。给定任意策略剖面 $\sigma = (\sigma_1, \dots, \sigma_n) \in \Sigma$, 其中 σ_{-i} 表示除玩家 i 之外所有其他玩家的策略, a_{-i} 表示除玩家 i 之外所有其他玩家的行动。对于玩家 i 和其纯策略 $a_i \in A_i$, 定义

$$u_i(a_i, \sigma_{-i}) = \sum_{a_{-i} \in A_{-i}} \left(\prod_{j \neq i} \sigma_j(a_j) \right) u_i(a_i, a_{-i}),$$

这个式子的含义是: 当其他玩家按照 σ_{-i} 进行混合时, 玩家 i 采取纯策略 a_i 所能获得的期望收益, 其中直积项表示出现情况 a_{-i} 的概率。类似地, 我们可以定义玩家 i 在混合策略 σ_i 下的期望收益:

$$u_i(\sigma_i, \sigma_{-i}) = \sum_{a_i \in A_i} \sigma_i(a_i)u_i(a_i, \sigma_{-i}).$$

观察这两个表达式, 它们都是由有限次乘积与有限次求和构成的多重线性函数, 因此在整个策略空间 Σ 上处处连续。这个连续性在后续的证明中至关重要。

有了期望收益的定义, 我们可以衡量某个纯策略相对于当前混合策略的改进程度。

基于收益差异，定义改进量

$$F_{ia_i}(\sigma) := u_i(a_i, \sigma_{-i}) - u_i(\sigma_i, \sigma_{-i}),$$

这个量衡量的是：如果玩家 i 从当前的混合策略 σ_i 改为纯策略 a_i ，其期望收益会增加多少（或减少多少）。为了只关注那些能带来改进的纯策略，我们进一步定义

$$k_{ia_i}(\sigma) := \max\{F_{ia_i}(\sigma), 0\}.$$

这里 $k_{ia_i}(\sigma)$ 的含义是：在策略剖面 σ 下，纯策略 a_i 相对于当前混合策略 σ_i 的非负改进额。如果该纯策略不能改进收益，改进量就是 0；如果能改进，改进量就等于收益的增加量。由于 F_{ia_i} 是连续函数，而函数 $x \mapsto \max\{x, 0\}$ 在 \mathbb{R} 上也是连续的，它们的复合 k_{ia_i} 在 Σ 上也必然是连续的。

为了构造我们需要的映射，我们还需要一个归一化因子。定义

$$S_i(\sigma) := 1 + \sum_{a'_i \in A_i} k_{ia'_i}(\sigma).$$

这个 $S_i(\sigma)$ 是所有改进量的总和再加上 1，它是连续函数且处处大于等于 1，因此处处非零。这个性质保证了我们后面定义的映射不会出现除以零的情况。

现在可以构造关键的映射了。对每个玩家 i ，定义映射

$$f_i(\sigma)(a_i) = \frac{\sigma_i(a_i) + k_{ia_i}(\sigma)}{S_i(\sigma)}.$$

这个映射的直观含义是：如果某个纯策略 a_i 能够改进当前收益（即 $k_{ia_i}(\sigma) > 0$ ），那么我们就增加它在混合策略中的权重；同时，为了保持概率分布的性质（总和为 1），我们用 $S_i(\sigma)$ 进行归一化。该式的分子与分母均为连续函数，且分母处处非零，因此 f_i 在 Σ 上连续。同时容易验证

$$\sum_{a_i \in A_i} f_i(\sigma)(a_i) = \frac{\sum_{a_i \in A_i} \sigma_i(a_i) + \sum_{a_i \in A_i} k_{ia_i}(\sigma)}{S_i(\sigma)} = \frac{1 + \sum_{a_i \in A_i} k_{ia_i}(\sigma)}{S_i(\sigma)} = 1,$$

故 $f_i(\sigma) \in \Delta(A_i)$ 。

把所有玩家的映射 f_i 组合起来，得到连续映射

$$f = (f_1, \dots, f_n) : \Sigma \rightarrow \Sigma.$$

让我们回顾一下这个映射的连续性是如何得到保证的：首先，多重线性表达保证了 F_{ia_i} 的连续性；然后，连续算子 $\max(\cdot, 0)$ 保持了 k_{ia_i} 的连续性；最后，有限和、加一操作以及处处非零的分母保证了做除法的连续性，从而每个分量 f_i 都连续，整体映射 f 也就连续了。

现在我们可以应用 Brouwer 不动点定理了。由于 Σ 是紧致且凸的集合， $f : \Sigma \rightarrow \Sigma$ 是连续映射，根据 Brouwer 不动点定理，必然存在 $\sigma^* \in \Sigma$ 满足 $f(\sigma^*) = \sigma^*$ 。

下面的关键步骤是证明这个不动点 σ^* 恰好就是 Nash 均衡。我们需要把不动点的性质逐步转化为 Nash 均衡的条件。写出分量形式的不动点关系，对于任意玩家 i 和任意纯策略 $a_i \in A_i$ ，有

$$\sigma_i^*(a_i) = \frac{\sigma_i^*(a_i) + k_{ia_i}(\sigma^*)}{1 + \sum_{a'_i} k_{ia'_i}(\sigma^*)}.$$

两边同乘分母并移项，得到恒等式

$$\sigma_i^*(a_i) \sum_{a'_i} k_{ia'_i}(\sigma^*) = k_{ia_i}(\sigma^*).$$

这个等式对所有玩家 i 和所有纯策略 a_i 都成立，它包含了关于不动点的重要信息。

接下来我们用反证法来推导 Nash 均衡条件。注意到 σ_i^* 是一个概率分布，因此必然存在至少一个动作 a'_i 使得 $\sigma_i^*(a'_i) > 0$ 。考虑所有具有正概率的动作，假设对每一个满足 $\sigma_i^*(a'_i) > 0$ 的 a'_i 都有严格不等式

$$u_i(a'_i, \sigma_{-i}^*) > u_i(\sigma_i^*, \sigma_{-i}^*),$$

这意味着所有被实际使用的纯策略（即概率为正的策略）都能严格改进当前的期望收益。将这些严格不等式按权重 $\sigma_i^*(a'_i)$ 求和，我们会得到

$$u_i(\sigma_i^*, \sigma_{-i}^*) = \sum_{a_i \in A_i} \sigma_i^*(a_i) u_i(a_i, \sigma_{-i}^*) = \sum_{\substack{a'_i \in A_i \\ \sigma_i^*(a'_i) > 0}} \sigma_i^*(a'_i) u_i(a'_i, \sigma_{-i}^*) > u_i(\sigma_i^*, \sigma_{-i}^*),$$

这显然是矛盾的——一个数不可能严格大于它自己。

因此，在正概率动作中必然存在某一动作 a''_i 满足

$$u_i(a''_i, \sigma_{-i}^*) \leq u_i(\sigma_i^*, \sigma_{-i}^*).$$

根据 k_{ia_i} 的定义，这意味着 $k_{ia''_i}(\sigma^*) = 0$ 。将此代入前面得到的恒等式，并利用 $\sigma_i^*(a''_i) > 0$ ，得到

$$\sigma_i^*(a''_i) \sum_{a'_i} k_{ia'_i}(\sigma^*) = 0 \implies \sum_{a'_i} k_{ia'_i}(\sigma^*) = 0.$$

这里关键的推理是：由于 $\sigma_i^*(a''_i) > 0$ ，而左边的乘积等于 0，所以求和项本身必须为 0。

既然各个 $k_{ia'_i}(\sigma^*) \geq 0$ 且其和为零，就必有 $k_{ia'_i}(\sigma^*) = 0$ 对所有 $a'_i \in A_i$ 成立。回到 k_{ia_i} 的定义，这等价于

$$u_i(a'_i, \sigma_{-i}^*) \leq u_i(\sigma_i^*, \sigma_{-i}^*) \quad \text{对任意 } a'_i \in A_i.$$

这正是说：玩家 i 无法通过改为任何一个纯策略来提高其期望收益。换句话说，当前的混合策略 σ_i^* 已经是对其他玩家策略 σ_{-i}^* 的最优反应。

由于 i 是任意选择的玩家，上述结论对所有玩家都成立。因此 σ^* 满足 Nash 均衡的定义：没有任何玩家能够通过单方面改变自己的策略来提高收益。根据引理，我们得到的不动点 σ^* 就是博弈的一个混合策略 Nash 均衡。命题得证。