

DAVENPORT · 圆法 · 高中详解版

三次型：双线性方程

Cubic forms: bilinear equations

本章要解决什么

我们正在研究一个 **三次型**（三次齐次多项式） $C(\mathbf{x})$ ，目标是证明：当变量个数 n 足够大时，方程 $C(\mathbf{x}) = 0$ 一定有非零整数解（也就是说， C “表示零”）。上一章（第 13 章）用圆法的技术，把问题逼到了一个岔路口——**引理 13.4** 说：要么出现“备择情形 A”，要么出现“备择情形 B”。备择情形 B 留到后面几章用圆法处理；**本章专门收拾备择情形 A**。

备择情形 A 是一个看似“坏”的情形：有**异常多**的整点对 $(\mathbf{x}, \mathbf{y})$ 满足一组叫做“双线性方程”的等式 $B_j(\mathbf{x} | \mathbf{y}) = 0$ 。本章要证明：这种“解多到反常”的情形，**恰恰**逼出我们想要的结论——存在非零整点 $\mathbf{z}$ 使 $C(\mathbf{z}) = 0$ 。于是无论走哪条岔路，结论都成立。

读完本章你会掌握：行列式与 Cramer 法则、矩阵的**秩与零空间维数**、**偏导数**、**≪ 记号**、把多项式方程组看成**代数簇**并用其**维数**数格点——以及如何把这些工具拼成一条完整的反证链。本章只用到代数与几何，**不出现**圆法里的积分和复指数，是全书里少见的“纯代数”一章。

阅读方式：黑字与公式是 Davenport 原书的内容（按译文逐段推进）；彩色框（目标 / 例 / 分步推演 / 符号卡）和插图是为高中生补的详解。凡原文写“显然”“立即可得”的地方，本页都补成一步一句的完整推导。

第 0 节 先把“舞台”和记号讲清楚

本章原文第一句就用到一堆上一章的记号。为了让本页真正“自足”，我们先花一节，把所有要用的对象从零讲一遍。看不懂第 13 章没关系，下面这些就够用了。

0.1 三次型 $C(\mathbf{x})$ 与对称系数 c_{ijk}

符号卡 · $\mathbf{x}$ 与 $C(\mathbf{x})$

$\mathbf{x} = (x_1, x_2, \dots, x_n)$ 是一个有 n 个分量的**向量**（黑体表示“一串数”，不是一个数）。本章里 $\mathbf{x}$ 的分量都取**整数**，这样的点叫**整点**。

$C(\mathbf{x})$ 是一个**三次型** (cubic form)，意思是“每一项都是三个变量相乘”的齐次三次多项式：

$$C(\mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n c_{ijk} x_i x_j x_k.$$

读法：三个求和号 $\sum_{i=1}^n$ 叠在一起，表示让下标 i, j, k 各自独立地从 1 跑到 n ，把所有 $c_{ijk} x_i x_j x_k$ 全加起来。每一项都是三个 x 相乘，所以是“三次”；不带常数项、不带一次二次项，所以是“齐次”。

符号卡·系数 c_{ijk} 与“对称”

c_{ijk} 是一串**整数**系数，每选定一组下标 (i, j, k) 就对应一个整数。我们假设它对三个下标**完全对称**：交换任意两个下标，值不变，例如

$$c_{ijk} = c_{jik} = c_{ikj} = c_{kji} = \dots$$

为什么能这样假设？因为在 $C(\mathbf{x})$ 里， $x_i x_j x_k$ 这个乘积本身就分不清谁先谁后（乘法可交换），所以把对应同一乘积的几个系数“平均一下”不改变 C ，却能让系数对称。对称性会在后面反复用到，务必记牢。

小例子（帮助理解 C 、 c_{ijk} ）

取 $n = 2$ ，设 $C(x_1, x_2) = x_1^3 + x_2^3$ 。展开看系数： x_1^3 来自 $(i, j, k) = (1, 1, 1)$ ，故 $c_{111} = 1$ ； x_2^3 来自 $(2, 2, 2)$ ，故 $c_{222} = 1$ ；其它系数全为 0。这就是“用 c_{ijk} 编码一个三次型”的含义。它“表示零”吗？要找非零整点使 $x_1^3 + x_2^3 = 0$ ，即 $x_1 = -x_2$ ，例如 $(1, -1)$ ，确实 $1 + (-1) = 0$ 。所以这个 C 表示零。

符号卡·“ C 表示零” (represents zero)

这是本章的核心目标词。它的意思是：**存在一个非零整点 $\mathbf{z} \neq \mathbf{0}$ ，使 $C(\mathbf{z}) = 0$** 。注意一定要“非零”，因为 $C(\mathbf{0}) = 0$ 是白送的（所有项都含 x ，代入全 0 当然得 0），没有意义。整章的反证法都建立在它的反面上：“ C 不表示零”就是“对一切整点 $\mathbf{x} \neq \mathbf{0}$ 都有 $C(\mathbf{x}) \neq 0$ ”。

0.2 双线性型 $B_j(\mathbf{x} \mid \mathbf{y})$

符号卡 · $B_j(\mathbf{x} | \mathbf{y})$

对两个向量 $\mathbf{x}, \mathbf{y}$, 定义 n 个**双线性型** (每个 $j = 1, \dots, n$ 一个):

$$B_j(\mathbf{x} | \mathbf{y}) = \sum_{i=1}^n \sum_{k=1}^n c_{ijk} x_i y_k.$$

“双线性”是说: 把 $\mathbf{y}$ 固定, 它是 $\mathbf{x}$ 的一次 (线性) 函数; 把 $\mathbf{x}$ 固定, 它又是 $\mathbf{y}$ 的一次函数。两边都“一次”, 合起来叫“双线性”。竖线“|”只是分隔 $\mathbf{x}$ 和 $\mathbf{y}$, 没有别的含义。

这几个 B_j 和 C 有两条关系, 本章证明里要反复用, 先在这里一次性算清楚。

1 B 对 $\mathbf{x}, \mathbf{y}$ 是对称的: $B_j(\mathbf{x} | \mathbf{y}) = B_j(\mathbf{y} | \mathbf{x})$ 。推导: $B_j(\mathbf{y} | \mathbf{x}) = \sum_{i,k} c_{ijk} y_i x_k$ 。把求和的哑标改名 $i \leftrightarrow k$, 得 $\sum_{i,k} c_{kji} y_k x_i$; 再用对称性 $c_{kji} = c_{ijk}$, 就得到 $\sum_{i,k} c_{ijk} x_i y_k = B_j(\mathbf{x} | \mathbf{y})$ 。

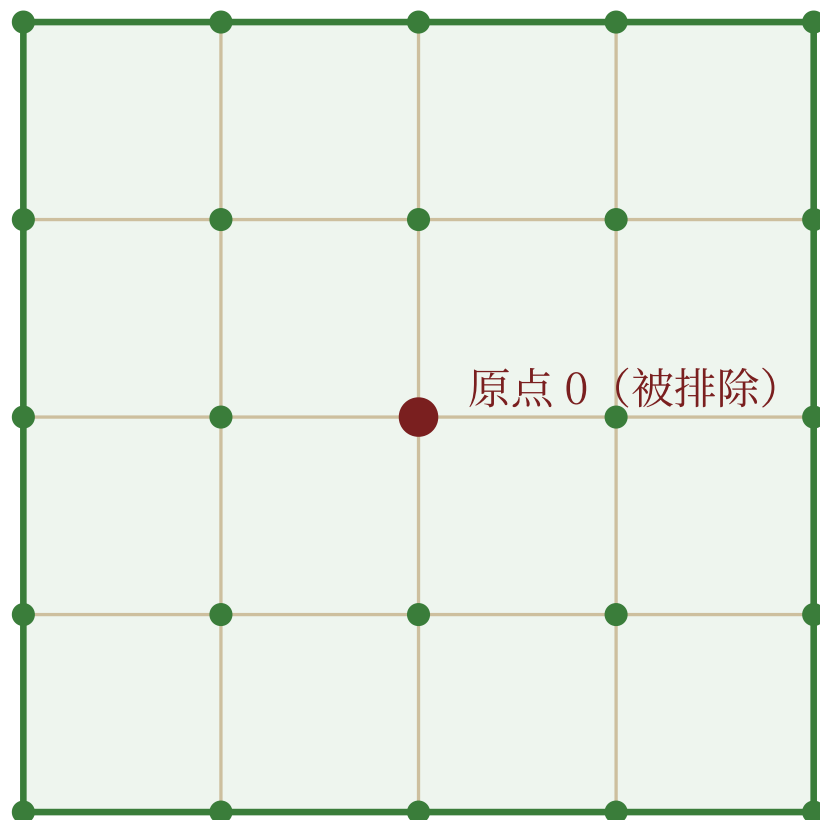
2 把中间下标当“编号”的换位公式: $\sum_{i,j} c_{ijk} x_i y_j = B_k(\mathbf{x} | \mathbf{y})$ 。推导: 对称性允许交换 c 的后两个下标 $c_{ijk} = c_{ikj}$, 于是 $\sum_{i,j} c_{ijk} x_i y_j = \sum_{i,j} c_{ikj} x_i y_j$; 对照 B_k 的定义 (x 配第一个下标、 y 配第三个下标、第二个下标是编号 k), 这正是 $B_k(\mathbf{x} | \mathbf{y})$ 。这条公式后面“求导后乘 y_j 求和”时是关键。

3 C 由 B 拼出: $\sum_j x_j B_j(\mathbf{x} | \mathbf{x}) = \sum_{i,j,k} c_{ijk} x_i x_j x_k = C(\mathbf{x})$ 。这解释了“双线性型”为什么是研究三次型的天然工具: 把 $\mathbf{y}$ 换成 $\mathbf{x}$ 、再用 $\mathbf{x}$ 加权求和, 就还原成 C 。

0.3 “大小” $|\mathbf{x}|$ 与“整点对很多”的含义

符号卡 · $|\mathbf{x}|$ 与 $0 < |\mathbf{x}|$

$|\mathbf{x}|$ 指各分量绝对值里最大的那个: $|\mathbf{x}| = \max(|x_1|, \dots, |x_n|)$ (这叫“最大模”或“盒子范数”)。于是 $|\mathbf{x}|$ 立方体盒子里; 再加上 $0 < |\mathbf{x}|$ 就是“非零且在盒子里”。这样的整点大约有 $(2R)^n \sim R^n$ 个——这个“ R^n 量级”是后面所有计数的基准尺子。



边长 $2R$ 的盒子，约 R^n 个整点

$$\{0 < |\mathbf{x}| \leq R\}$$

符号卡 · $\ll$ 、 $O(\cdot)$ 与 ε

$A \ll B$ (读作 “ A 远小于等于 B ”，与 $A = O(B)$ 同义) 表示：存在一个不依赖于 R 的常数 c ，使得 $|A| \leq cB$ 对所有大的 R 成立。它只关心 “增长的量级”，把具体常数藏起来。例如 “盒子里的格点数 $\ll R^n$ ” 就是说点数不超过某个常数乘 R^n 。

ε (epsilon) 是一个 “任意小的固定正数”。写 $R^{n+\varepsilon}$ 是为了表达 “几乎是 R^n 、但允许多一丁点”。本章里它像一个可调旋钮：证明时把它取小，结论里允许它出现，最后两边的 ε 一对，矛盾就成立 (见第 6 节，把本章的 ε 取成情形 A 那个 ε 的一半)。

第 1 节 把双线性方程组看成线性方程组，引出 Hessian $H(\mathbf{x})$

(对应原文开头到公式 (14.2) 之间的五个自然段。)

原文第 1 段大意：本章研究引理 13.4 的“备择情形 A”——对某个任意大的 R ，存在多于 $R^{n+\varepsilon}$ 对整点 $\mathbf{x}, \mathbf{y}$ 满足 $0 < |\mathbf{x}|$

$$B_j(\mathbf{x} | \mathbf{y}) = 0, \quad (1 \leq j \leq n). \quad (14.1)$$

逐句拆解这句话。它在说一个“坏消息”：满足那 n 个等式 (14.1) 的整点对，多得反常。“反常”在哪？所有整点对一共才约 R^{2n} 个 ($\mathbf{x}$ 有 R^n 种、 $\mathbf{y}$ 有 R^n 种)。而 (14.1) 是 n 个方程，凭直觉每加一个方程会把解“砍掉一个维度”，本该只剩约 R^{2n} 经过 n 个方程后降到 $R^{2n-n} = R^n$ 量级。情形 A 说的却是“多于 $R^{n+\varepsilon}$ ”，比这个直觉值还多出一个 R^ε 因子——这点点“超额”，正是本章要利用的破绽。

原文第 2 段大意：按上一章末尾的约定，已经排除了 $\mathbf{x} = \mathbf{0}$ 与 $\mathbf{y} = \mathbf{0}$ （所以才写 $0 < |\mathbf{x}|$ 、 $0 < |\mathbf{y}|$ ）。本章是自足的，将证明：上述假设蕴涵存在非零整点 $\mathbf{z} \neq \mathbf{0}$ 使 $C(\mathbf{z}) = 0$ ，即 C 表示零。原文还说：其实更弱的假设就够——只要点对个数“多于 AR^n ”、其中 A 大于 n 的某个函数即可；但用 ε 写起来更干净。

为什么用 $R^{n+\varepsilon}$ 而不用 AR^n 这是一个“取舍”说明。两种写法证的是同一件事，但 $R^{n+\varepsilon}$ （“比 R^n 多一个幂次”）在做不等式估计时只需比较指数 $n + \varepsilon$ 与 n ，一眼就分胜负；而 AR^n （“同一幂次、但常数特别大”）要去追踪那个常数 A 到底多大、和 n 的函数怎么比，叙述繁琐。作者选了简洁的写法——这不改变数学，只改变行文。

原文第 3 段大意：把某个特定的 $\mathbf{x}$ 固定下来，那么 (14.1) 就成了关于未知数 $\mathbf{y} = (y_1, \dots, y_n)$ 的 n 个线性方程，它们的行列式（系数矩阵的行列式）记作 $H(\mathbf{x})$ 。

这一步是全章的“视角转换”，必须讲透。把 $\mathbf{x}$ 固定后，看 $B_j(\mathbf{x} | \mathbf{y}) = \sum_k \left(\sum_i c_{ijk} x_i \right) y_k$ 。括号里的 $\sum_i c_{ijk} x_i$ 此刻是一个已知数（因为 $\mathbf{x}$ 定了），把它记成矩阵元

$$M_{jk}(\mathbf{x}) = \sum_{i=1}^n c_{ijk} x_i.$$

于是 (14.1) 就是 $\sum_k M_{jk}(\mathbf{x}) y_k = 0$ ，写成矩阵形式就是齐次线性方程组

$$M(\mathbf{x}) \mathbf{y} = \mathbf{0}.$$

符号卡·行列式 $\det$ 与“齐次线性方程组何时有非零解”

一个 $n \times n$ 方阵 M 的行列式 $\det M$ 是一个由它的元素算出来的数（ 2×2 时 $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$ ）。高中只需记住它的一条关键性质：

$$\text{齐次方程组 } M\mathbf{y} = \mathbf{0} \text{ 有非零解} \iff \det M = 0.$$

直观理由：若 $\det M \neq 0$ ，则 M 可逆，两边左乘 M^{-1} 得 $\mathbf{y} = \mathbf{0}$ ，只有零解；只有当 $\det M = 0$ （矩阵“退化、压扁了维度”）时，才会冒出非零解。

因此“是否存在非零的 $\mathbf{y}$ 满足 (14.1)”这件事，完全由一个数 $H(\mathbf{x}) = \det M(\mathbf{x})$ 是否为零来决定：

$$H(\mathbf{x}) = \det \left(\sum_{i=1}^n c_{ijk} x_i \right), \quad (1 \leq j, k \leq n). \quad (14.2)$$

读法：行列式里第 j 行第 k 列的元素是 $\sum_i c_{ijk} x_i$ ，让 j, k 各跑 1 到 n 就排成一个 $n \times n$ 矩阵，取它的行列式即得 $H(\mathbf{x})$ 。

固定 $\mathbf{x}$

$$\begin{aligned} B_j(\mathbf{x} | \mathbf{y}) &= 0 \\ \text{关于 } \mathbf{y} \text{ 的} \\ n \text{ 个线性方程} \\ M(\mathbf{x}) \mathbf{y} &= 0 \end{aligned}$$



$$\begin{aligned} \text{有非零解 } \mathbf{y} ? \\ \Leftrightarrow H(\mathbf{x}) = \det M(\mathbf{x}) = 0 \\ \text{把“解的存在性”} \\ \text{压缩成一个数是否为 0} \end{aligned}$$

核心视角：固定 $\mathbf{x}$ ，双线性方程组退化为线性方程组；“有没有非零 $\mathbf{y}$ ”等价于“行列式 $H(\mathbf{x})$ 是不是 0”。

原文第 4–5 段大意： $H(\mathbf{x})$ 就是三次型 $C(\mathbf{x})$ 的 **Hessian**（黑塞行列式）。它是关于 $\mathbf{x}$ 的一个 n 次型；更准确说是“表观次数（apparent degree）为 n ”，因为它有可能恒等于零（即作为多项式整个是 0）。所以第一件要做的事，是证明：只要 $C(\mathbf{x})$ 不表示零， $H(\mathbf{x})$ 就不会恒为零。

符号卡·Hessian（黑塞）与“次数”

一个函数 C 的 **Hessian 矩阵** 是它所有二阶偏导数排成的方阵，元素是 $\frac{\partial^2 C}{\partial x_j \partial x_k}$ 。对我们的 C 算一下（用系数全对称）：先 $\frac{\partial C}{\partial x_j} = 3 \sum_{i,k} c_{ijk} x_i x_k$ ，再 $\frac{\partial^2 C}{\partial x_j \partial x_k} = 6 \sum_i c_{ijk} x_i = 6 M_{jk}(\mathbf{x})$ 。可见 Hessian 矩阵恰是 $6M(\mathbf{x})$ ，其行列式 $= 6^n H(\mathbf{x})$ 。差一个常数 6^n 无关紧要，所以把 $H(\mathbf{x})$ 叫 Hessian。

为什么次数是 n ? 矩阵每个元素 $\sum_i c_{ijk}x_i$ 是 $\mathbf{x}$ 的一次式。行列式是把每行各取一个元素相乘 (n 个一次式相乘是 n 次) 再求和, 所以 $H(\mathbf{x})$ 是 $\mathbf{x}$ 的 n 次齐次多项式——除非这些 n 次项彼此抵消, 使整个多项式变成 0。“表观次数”就是提醒: 表面看是 n 次, 但实际可能恒为 0, 那时“次数”说法就名存实亡了。这正是引理 14.1 要排除的退化。

第 2 节 引理 14.1: C 不表示零 $\implies$ Hessian 不恒为零

引理 14.1. 若对所有非零整点 $\mathbf{x} \neq \mathbf{0}$ 都有 $C(\mathbf{x}) \neq 0$, 则 $H(\mathbf{x})$ 不恒等于零。

这条引理在干嘛 它是整章的“奠基石”: 保证我们手上的 Hessian $H(\mathbf{x})$ 是个货真价实的 n 次型 (不是恒零的废物), 这样它的零点集 $H(\mathbf{x}) = 0$ 才是一张真正的“ $n - 1$ 维曲面”, 上面的整点才稀少 ($\ll R^{n-1}$ 个)。证明用反证法: 假设 $H \equiv 0$, 构造出一个非零的 $\mathbf{y}$ 使 $C(\mathbf{y}) = 0$, 与“ C 不表示零”冲突。

下面把原文的证明逐句补成完整推导。原文有几处“一笔带过”, 我们停下来算清楚。

2.1 反设与“恒等秩”的设定

原文: 假设 $H(\mathbf{x}) \equiv 0$ 恒成立。设 $n - r$ ($r \geq 1$) 为 $H(\mathbf{x})$ 的恒等秩; 即假设 (14.2) 中矩阵的所有 $n - r + 1$ 阶子行列式都关于 $\mathbf{x}$ 恒为零, 但存在某个 $n - r$ 阶子行列式不恒为零。为方便, 设这个 $n - r$ 阶行列式 (记作 Δ) 位于左上角。

符号卡 · 矩阵的秩、子行列式、恒等秩

子行列式 (子式、minor): 从矩阵里挑出某 m 行、某 m 列, 交点处的元素组成一个 $m \times m$ 小方阵, 它的行列式就叫一个“ m 阶子行列式”。

秩 (rank): 使得“存在不为零的 m 阶子行列式”的最大 m 。秩 $= m$ 直观上表示“矩阵真正独立的行 (或列) 有 m 个”。

恒等秩: 这里矩阵元素是 $\mathbf{x}$ 的多项式, 所以子行列式也是 $\mathbf{x}$ 的多项式。“某子式恒为零”指它作为多项式是 0。把秩的定义套到“恒等”上: 恒等秩 $= n - r$ 意思是——有一个 $n - r$ 阶子式不恒为零, 而所有 $n - r + 1$ 阶子式都恒为零。因为我们已反设 $H = \det M \equiv 0$ (这是 n 阶子式), 所以恒等秩 $\leq n - 1$, 即 $r \geq 1$ 。

“设 Δ 在左上角”是可以这么假设的: 调换行的顺序、列的顺序, 只会让行列式变号, 不影响“是否为零”。所以总能把那个不恒为零的 $n - r$ 阶子式搬到左上角, 记

$$\Delta = \det (M_{jk}(\mathbf{x}))_{1 \leq j, k \leq n-r} (\neq 0).$$

原文：那么 (14.1) 的前 $n-r$ 个方程就蕴涵其余所有方程，因为后面诸行与前 $n-r$ 行线性相关。

1 为什么“后面的行依赖前面的行”？把矩阵 $M(\mathbf{x})$ 的第 j 行记作

$\mathbf{m}_j = (M_{j1}, \dots, M_{jn})$ 。恒等秩是 $n-r$ ，意味着任取 $n-r+1$ 行都线性相关（否则会有不恒零的 $n-r+1$ 阶子式）。前 $n-r$ 行（在前 $n-r$ 列上）的行列式 $\Delta \neq 0$ ，说明这 $n-r$ 行线性无关。于是再添任何一行 $\mathbf{m}_t$ ($t > n-r$)，这 $n-r+1$ 行必相关，而前 $n-r$ 行已无关，所以多出来的 $\mathbf{m}_t$ 只能是前 $n-r$ 行的线性组合。

2 “蕴涵”的含义：既然第 t 行是前面行的组合，那么方程 $\mathbf{m}_t \cdot \mathbf{y} = 0$ 也是前 $n-r$ 个方程 $\mathbf{m}_1 \cdot \mathbf{y} = 0, \dots$ 的同样组合。只要前 $n-r$ 个成立，第 t 个自动成立。所以解 (14.1) 只需解前 $n-r$ 个方程。

2.2 用 Cramer 法则写出一个非零解

原文：设 Δ_j 表示把 Δ 的第 j 列换成第 $(n-r+1)$ 列后得到的行列式。那么 $B_j(\mathbf{x} | \mathbf{y}) = 0$ 的一个特解是

$$y_1 = \Delta_1, \dots, y_{n-r} = \Delta_{n-r}, y_{n-r+1} = -\Delta, y_{n-r+2} = \dots = 0.$$

（这由 Cramer 法则得出：求解前 $n-r$ 个方程，令 $y_{n-r+1} = 1, y_{n-r+2} = 0, \dots$ ，再整体乘以 $-\Delta$ 。）

符号卡 · Cramer 法则

解 m 元线性方程组 $\mathbf{A}\mathbf{u} = \mathbf{b}$ (A 是 $m \times m$ 、 $\det A \neq 0$) 时，Cramer 法则说：第 k 个未知数

$$u_k = \frac{\det A^{(k)}}{\det A},$$

其中 $A^{(k)}$ 是把 A 的第 k 列换成右端 $\mathbf{b}$ 后的矩阵。这是高中行列式知识的标准内容，记住“换列、求商”即可。

把它套到我们的情形，一步一步来：

1 只保留前 $n-r$ 个方程，并把 $y_{n-r+2}, \dots, y_n$ 全设为 0，再让 $y_{n-r+1} = 1$ 。这样前 $n-r$ 个方程 $\sum_{k=1}^n M_{jk} y_k = 0$ 变成只含 $y_1, \dots, y_{n-r}$ 的方程：

$$\sum_{k=1}^{n-r} M_{jk} y_k = -M_{j, n-r+1}, \quad (j = 1, \dots, n-r).$$

右端就是“第 $(n-r+1)$ 列、取负号”。

- 2 这是一个 $(n-r) \times (n-r)$ 方程组，系数矩阵行列式正是 $\Delta \neq 0$ ，可用 Cramer 法则。第 k 个解

$$y_k = \frac{\det(\text{把 } \Delta \text{ 的第 } k \text{ 列换成右端})}{\Delta}.$$

右端是 $-(\text{第 } n-r+1 \text{ 列})$ ，那个分子行列式就是 $-\Delta_k$ （负号可提出来）。故

$$y_k = -\Delta_k / \Delta.$$

- 3 分母 Δ 是个多项式，做除法不方便。把所有 y 整体乘以 $-\Delta$ （齐次线性方程的解乘以同一非零量仍是解）。于是

$$y_k \rightsquigarrow (-\Delta) \cdot \frac{-\Delta_k}{\Delta} = \Delta_k \quad (k \leq n-r), \quad y_{n-r+1} \rightsquigarrow (-\Delta) \cdot 1 = -\Delta,$$

其余仍为 0。这就得到原文写的特解。

- 4 **关键收获：**每个 y_k 现在是 $\mathbf{x}$ 的整系数多项式（因为 Δ, Δ_k 都是由整数 c_{ijk} 组成的行列式，且行列式只用加减乘）。而且 $y_1, \dots, y_n$ 不全恒为零，因为其中 $y_{n-r+1} = -\Delta \not\equiv 0$ 。这一点马上要用。

2.3 对恒等式求导，挤出 $C(\mathbf{y}) = 0$

现在 $\mathbf{y} = \mathbf{y}(\mathbf{x})$ 是 $\mathbf{x}$ 的多项式向量，且对每个 $\mathbf{x}$ 满足 (14.1)。也就是说，作为 $\mathbf{x}$ 的多项式恒等式，

$$B_j(\mathbf{x} \mid \mathbf{y}) = \sum_{i,k} c_{ijk} x_i y_k = 0 \quad (\text{对所有 } j, \text{ 关于 } x_1, \dots, x_n \text{ 恒成立}).$$

符号卡·偏导数 $\partial/\partial x_\nu$

把多变量函数对某一个变量 x_ν 求导，其余变量当常数，就是偏导数，记作 $\frac{\partial}{\partial x_\nu}$ 。规则和高中一元求导一样，只是只动 x_ν 。两条要用的事实：(i) $\frac{\partial x_i}{\partial x_\nu} = \begin{cases} 1 & i = \nu \\ 0 & i \neq \nu \end{cases}$ （互不相同的变量彼此无关）；(ii) 对“恒等于 0 的多项式”求导，结果仍恒为 0（0 的导数是 0）。我们对一个恒等式两边同时求导，等号继续成立。

对上面的恒等式关于 x_ν 求导（用乘积法则；注意 y_k 也含 x ）：

$$\sum_k c_{\nu j k} y_k + \sum_{i,k} c_{i j k} x_i \frac{\partial y_k}{\partial x_\nu} = 0 \quad (\text{对所有 } j, \nu). \quad (*)$$

1 第一项怎么来的：对 $\sum_{i,k} c_{i j k} x_i y_k$ 中的 x_i 求导，只有 $i = \nu$ 那项存活， $\partial x_i / \partial x_\nu = 1$ ，得到 $\sum_k c_{\nu j k} y_k$ 。

2 第二项怎么来的：再对 y_k 求导（ x_i 当系数留住），得 $\sum_{i,k} c_{i j k} x_i \partial y_k / \partial x_\nu$ 。两项之和为 0，因为左边恒为 0。

原文：乘以 y_j 并对 j 求和，并注意到 $\sum_{i,j} c_{i j k} x_i y_j = 0$ 对所有 k 成立，得到 $\sum_{j,k} c_{\nu j k} y_j y_k = 0$ 对所有 ν 成立。

1 把 (*) 两边乘 y_j ，再对 j 从 1 到 n 求和：

$$\sum_{j,k} c_{\nu j k} y_j y_k + \sum_j \left(\sum_{i,k} c_{i j k} x_i \frac{\partial y_k}{\partial x_\nu} \right) y_j = 0.$$

2 看第二个和。把它按 k 归并： $\sum_k \left(\sum_{i,j} c_{i j k} x_i y_j \right) \frac{\partial y_k}{\partial x_\nu}$ 。括号里正是第 0.2 节的“换位公式” $\sum_{i,j} c_{i j k} x_i y_j = B_k(\mathbf{x} | \mathbf{y})$ 。

3 而 $\mathbf{y}$ 是 (14.1) 的解，所以 $B_k(\mathbf{x} | \mathbf{y}) = 0$ 对一切 k ！于是整个第二个和 $= \sum_k 0 \cdot \frac{\partial y_k}{\partial x_\nu} = 0$ ，凭空消失。

4 只剩 $\sum_{j,k} c_{\nu j k} y_j y_k = 0$ ，对所有 ν 成立。这是一组关于 $\mathbf{y}$ 的二次恒等式。

原文：这特别地蕴涵 $C(\mathbf{y}) = 0$ 。

1 把刚得到的 $\sum_{j,k} c_{\nu j k} y_j y_k = 0$ 两边乘 y_ν ，对 ν 求和：

$$\sum_\nu y_\nu \sum_{j,k} c_{\nu j k} y_j y_k = \sum_{\nu,j,k} c_{\nu j k} y_\nu y_j y_k.$$

2 右边正是 $C(\mathbf{y})$ 的定义（三个下标 ν, j, k 各跑 1 到 n ，乘 $y_\nu y_j y_k$ ）。而左边是 $\sum_\nu y_\nu \cdot 0 = 0$ 。所以 $C(\mathbf{y}) = 0$ （作为 $\mathbf{x}$ 的多项式恒成立）。

2.4 代入一个具体整点，导出矛盾

原文：取 $\mathbf{x}$ 为任意一个使 $\mathbf{y} \neq \mathbf{0}$ 的整点（这可能，因为 $\mathbf{y}$ 不恒为 $\mathbf{0}$ ），便与假设矛盾。引理得证。

1 $\mathbf{y}(\mathbf{x})$ 是不恒为零的整系数多项式向量（2.2 末尾已证 $y_{n-r+1} = -\Delta \neq 0$ ）。一个不恒为零的多项式，必能找到某个整数代入点使它非零（多项式的零点很“稀”，不可能在所有整点都为零）。取这样一个整点 $\mathbf{x}_0$ ，则 $\mathbf{y}_0 = \mathbf{y}(\mathbf{x}_0)$ 是一个非零整点。

2 由 2.3, $C(\mathbf{y}) = 0$ 对所有 $\mathbf{x}$ 成立，特别 $C(\mathbf{y}_0) = 0$ 。于是 $\mathbf{y}_0 \neq \mathbf{0}$ 却使 $C(\mathbf{y}_0) = 0$ 。

3 这与引理前提“ C 不表示零（对一切非零整点 $C \neq 0$ ）”直接冲突。矛盾来自最初的反设 $H \equiv 0$ 。故 $H(\mathbf{x})$ 不恒为零。■

回看证明的“魔术”在哪 整段证明的精彩之处：从“ $H \equiv 0$ ”这条关于行列式的假设出发，借 Cramer 法则把“退化”翻译成一个具体的解 $\mathbf{y}(\mathbf{x})$ ，再用求导 + 系数对称性把双线性等式 $B_j = 0$ 升级成二次等式，最终拼出三次式 $C(\mathbf{y}) = 0$ 。一路都是“代数恒等式变形”，没有任何分析或极限，干净利落。

第 3 节 从 $r = 1$ 到一般 r ：要数的到底是什么

原文段：上一引理表明，那些使 $B_j(\mathbf{x} \mid \mathbf{y}) = 0$ 关于 $\mathbf{y}$ 有非零解的点 $\mathbf{x}$ ，满足非恒等的方程 $H(\mathbf{x}) = 0$ 。于是满足 $\{0 < |\mathbf{x}| \leq \mathbf{x}\}$

这一段是路线图，要看懂三件事。

符号卡·线性无关 / 解空间维数 / 秩与零度

齐次方程组 $M(\mathbf{x})\mathbf{y} = \mathbf{0}$ 的全体解构成一个解空间。其中“ r 个线性无关解”指能找到 r 个解 $\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(r)}$ ，谁也不是其他几个的线性组合；最大能找到的这种个数叫解空间的维数，也叫矩阵的零度（nullity）。线性代数的秩-零度定理说：

$$\text{秩}(M) + \text{零度}(M) = n.$$

所以“有 r 个线性无关解”就等于“秩 $M(\mathbf{x}) = n - r$ ”。 r 越大、秩越小、矩阵越退化。

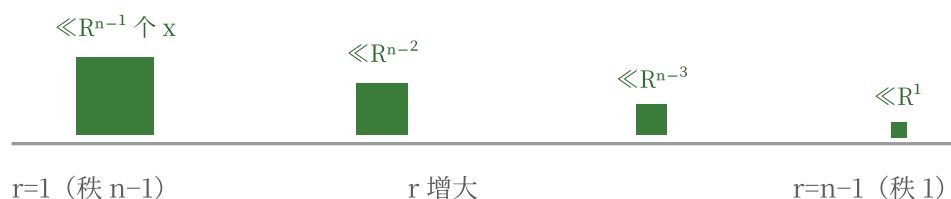
1 “有非零解 $\iff H(\mathbf{x}) = 0$ ”：由第 1 节，固定 $\mathbf{x}$ 时 $M(\mathbf{x})\mathbf{y} = \mathbf{0}$ 有非零解当且仅当 $\det M(\mathbf{x}) = H(\mathbf{x}) = 0$ 。引理 14.1 又保证 H 不恒为零，所以 $H(\mathbf{x}) = 0$ 是一

个“真方程”（不是 $0 = 0$ 那种废话方程）。

2 “这种 $\mathbf{x}$ 只有 $\ll R^{n-1}$ 个”： $H(\mathbf{x}) = 0$ 是一个不恒为零的 n 次多项式方程，它在 n 维空间里描出一张 $(n-1)$ 维曲面（少一维，因为多一个约束）。一张 $(n-1)$ 维曲面，在边长 $2R$ 的盒子里大约只含 $\ll R^{n-1}$ 个整点（比满盒子的 R^n 少一个 R 的因子）。这是“维数低一档、格点少一个幂次”的直观，第 4 节会把它讲严谨。

3 推广目标：“ $r = 1$ ”对应秩恰好掉到 $n-1$ （零度 1）。现在要对每个 r 都证：秩掉到 $n-r$ （零度 r ）的 $\mathbf{x}$ ，只有 $\ll R^{n-r}$ 个。为什么这是合理的？“秩掉得越多”是越苛刻的退化条件，要同时让更多子式为零，约束更多，满足的 $\mathbf{x}$ 应当更稀少，掉到 $n-r$ 维曲面上，格点 $\ll R^{n-r}$ 。这就是引理 14.3 的内容；为证它，先要一条代数几何的工具引理 14.2。

秩 $M(\mathbf{x})=n-r \Leftrightarrow$ 解空间维数（零度） $=r$



退化越深（ r 越大），单个 $\mathbf{x}$ 的解空间越大，但满足这种退化的 $\mathbf{x}$ 本身越稀少。一多一少，乘起来恰好被 R^n 量级卡住——这是第 6 节算总账的关键。

原文段：先处理一个初等代数几何的问题是方便的。——这是在为引理 14.2 铺垫：要把“低维曲面上格点少”这件直觉严格化，需要“代数簇的维数”这个语言。

第 4 节 引理 14.2：一条代数几何工具（“低维 $\Rightarrow$ 格点少 $\Rightarrow$ 雅可比矩阵退化”）

引理 14.2. 设 $f_1(\mathbf{x}), \dots, f_N(\mathbf{x})$ 是 $x_1, \dots, x_n$ 的整系数型，且 N 与它们的公共次数关于 n 有界。若对某个任意大的 R ，存在整点集 $\mathcal{X}$ ，其中 $\mathbf{x}$ 满足 $[\mathbf{x}] \leq R$ ，且 $f_i(\mathbf{x}) = 0$ 对 $i = 1, \dots, N$ 成立，则 $n \leq N$ 。

$0 < |\mathbf{x}|$

注. 这组关于 $\partial f_i / \partial x_\nu$ 的方程，等价于断言矩阵 $(\partial f_i / \partial x_\nu)$ ($i = 1, \dots, N$, $\nu = 1, \dots, n$) 的秩 $\leq r - 1$ 。这由初等矩阵理论得出。

先把结论“翻译成人话” 这个矩阵 $(\partial f_i / \partial x_\nu)$ 叫雅可比矩阵 (Jacobian)，第 i 行是函数 f_i 的梯度 (各偏导排一行)。引理说：如果方程组 $f_i = 0$ 的整数解“多到超过 $R^{n-r+\varepsilon}$ 个”，那么必有某个解点，在那里所有这些梯度只张成一个 $\leq r - 1$ 维的空间 (雅可比矩阵秩 $\leq r - 1$)。直觉：解太多 $\implies$ 解集是个相当高维的曲面 $\implies$ 在它上面，约束函数的梯度方向不够多 (互相挤在低维里)。下面的“秩 $\leq r - 1$ ”与“能写成 $\sum_{\rho=1}^{r-1} T_{i\rho} D_{\rho\nu}$ ”是同一回事——见下卡。

符号卡·为什么“秩 $\leq r-1$ ”就能写成那种求和

一个矩阵 $G_{i\nu}$ (这里 $G_{i\nu} = \partial f_i / \partial x_\nu$) 的秩 $\leq r - 1$ ，按定义等于：它的每个行向量都落在某个 $\leq r - 1$ 维的空间里，即存在 $r - 1$ 个“基向量” $D_{\rho\cdot} = (D_{\rho 1}, \dots, D_{\rho n})$ ($\rho = 1, \dots, r - 1$) 使每一行都是它们的线性组合：

$$G_{i\nu} = \sum_{\rho=1}^{r-1} T_{i\rho} D_{\rho\nu},$$

其中系数 $T_{i\rho}$ 是第 i 行的“配方比例”。注意 $D_{\rho\nu}$ 不带 i (基与行无关)， $T_{i\rho}$ 不带 ν (配方比例与列无关)——这正是结论里反复强调的“ D 与 f 无关、 T 与 ν 无关”。所以“写成那种求和” $\iff$ “秩 $\leq r - 1$ ”，就是这条注的意思。

4.1 把方程组的解集看成“代数簇”，挑出一块主分量

原文：方程 $f_1 = \dots = f_N = 0$ 在 n 维空间 (或 $(n - 1)$ 维射影空间) 中定义了一个代数簇。这样的簇可表示为若干绝对不可约簇之并，且在当前假设下这些不可约簇的个数关于 n 有界。因此存在其中之一，记作 $\mathcal{V}$ ，包含 $\mathcal{X}$ 中多于 $R^{n-r+\frac{1}{2}\varepsilon}$ 个点。

符号卡·代数簇 / 不可约 / 维数

代数簇 (variety)：一组多项式方程的公共解集，几何上就是若干曲面交在一起得到的图形。例如 $n = 3$ 里 $x_1 = 0$ 是一张平面， $x_1^2 + x_2^2 = 1$ 是一个柱面。

不可约 (irreducible)：一个簇若不能拆成两块更小的簇之并，就叫不可约，相当于“一整块、不是几片拼起来的”。任何代数簇都能唯一地拆成有限多块不可约簇的并 (像把一个图形拆成几个连通的整片)。“绝对”指在复数范围内不可约。

维数 s ：这块簇“几维”。一条曲线 $s = 1$ 、一张曲面 $s = 2$ ，依此类推。本节只需用到维数的一个操作性刻画（见下）。

1 把解集拆成有限多块不可约簇之并。块数“关于 n 有界”——因为 N 和次数都被 n 控制住了，分量数不会失控（这是代数几何里 Bézout 型的界）。设块数为 $\leq B(n)$ 。

2 $\mathcal{X}$ 的全部点（多于 $R^{n-r+\varepsilon}$ 个）分布在这 $\leq B(n)$ 块里。由**抽屉原理**：若每块都不超过 $R^{n-r+\frac{1}{2}\varepsilon}$ 个点，则总数 $\leq B(n)R^{n-r+\frac{1}{2}\varepsilon}$ ，当 R 充分大时这小于 $R^{n-r+\varepsilon}$ （因为多出的 $R^{\frac{1}{2}\varepsilon}$ 因子最终压过常数 $B(n)$ ），矛盾。故至少有一块 $\mathcal{V}$ 含多于 $R^{n-r+\frac{1}{2}\varepsilon}$ 个 $\mathcal{X}$ 中的点。

为什么 ε 要“减半” 把要求从 $R^{n-r+\varepsilon}$ 降到 $R^{n-r+\frac{1}{2}\varepsilon}$ ，是给抽屉原理“留出余量”：丢掉的那一半 $\frac{1}{2}\varepsilon$ 用来吸收常数因子 $B(n)$ 。这是解析数论里典型的“拿 ε 当缓冲”手法。

4.2 用维数把“格点数”与 s 挂钩，得 $s \geq n - r + 1$

原文：与 $\mathcal{V}$ 相联系的是其维数 s 。我们只需 s 的下述性质：不可约簇 $\mathcal{V}$ 可分解为有限多部分，在每部分上 $x_1, \dots, x_n$ 中有 s 个是独立变量、另 $n - s$ 个是它们的单值可微函数。由此 $\mathcal{V}$ 含 $\ll R^s$ 个满足 $\|(\mathbf{x})\|$

1 **维数 s 的操作性含义**：在 $\mathcal{V}$ 的局部，可以挑出 s 个坐标当“自由变量”（任意取值），剩下 $n - s$ 个坐标被它们唯一决定（是它们的可微函数）。这正说明 $\mathcal{V}$ 是“ s 维”的——有 s 个自由度。

2 **数 $\mathcal{V}$ 上的整点**：一个整点落在 $\mathcal{V}$ 上、且 $\|(\mathbf{x})\|$

3 **两边一比**：一方面 $\mathcal{V}$ 上的 $\mathcal{X}$ -点多于 $R^{n-r+\frac{1}{2}\varepsilon}$ ；另一方面 $\mathcal{V}$ 上的整点 $\ll R^s$ 。于是 $R^{n-r+\frac{1}{2}\varepsilon} \lesssim R^s$ 。当 R 大时，比较指数得 $s \geq n - r + \frac{1}{2}\varepsilon$ ；但 s 是整数、 $n - r$ 也是整数，能严格超过 $n - r$ 的整数最小就是 $n - r + 1$ 。故

$$s \geq n - r + 1, \quad \text{即} \quad n - s \leq r - 1.$$

最后这个 $n - s \leq r - 1$ 是后面定结论的关键不等式。

s 维簇 V 上的整点数 $\ll R^s$

自由坐标 1: $\ll R$ 种

自由坐标 2: $\ll R$ 种

$\vdots$ (共 s 个)



其余 $n-s$ 个坐标
被自由坐标唯一决定

“维数 s ” 的用法: s 个自由坐标各有 $\ll R$ 种取值, 相乘得 $\ll R^s$ 个整点。解多 $\implies s$ 大 $\implies n-s$ 小。

4.3 在 $\mathcal{V}$ 上, 凡处处为零的函数其梯度只占 $n-s$ 维

原文: 考虑 $\mathcal{X}$ 在 $\mathcal{V}$ 上任一点的邻域, 可设 $x_{s+1}, \dots, x_n$ 是 $x_1, \dots, x_s$ 的单值可微函数。设 $f(x_1, \dots, x_n)$ 是在 $\mathcal{V}$ 上处处为零的任意可微函数。对恒等式 $f=0$ 关于 x_ν ($\nu=1, \dots, s$) 微分, 得

$$f^{(\nu)} + f^{(s+1)} \frac{\partial x_{s+1}}{\partial x_\nu} + \dots + f^{(n)} \frac{\partial x_n}{\partial x_\nu} = 0,$$

其中 $f^{(j)} = \partial f / \partial x_j$ (把 $x_1, \dots, x_n$ 当独立变量算)。

1 局部图像: 取 $x_1, \dots, x_s$ 为自由变量, 则 $x_{s+1}, \dots, x_n$ 是它们的函数。于是在 $\mathcal{V}$ 上, f 实际上只依赖 $x_1, \dots, x_s$ (其余通过它们间接依赖), 且 $f \equiv 0$ 。

2 链式法则: 对 $f(x_1, \dots, x_s, x_{s+1}(\cdot), \dots, x_n(\cdot)) = 0$ 关于自由变量 x_ν ($\nu \leq s$) 求导。直接依赖给出 $f^{(\nu)}$; 通过 $x_{s+\rho}$ 的间接依赖各给出 $f^{(s+\rho)} \cdot \frac{\partial x_{s+\rho}}{\partial x_\nu}$ 。全部相加为 0 (因为 f 恒为 0)。这就是上式。

原文: 于是对 $\nu=1, \dots, s$:

$$f^{(\nu)} = \sum_{\rho=1}^{n-s} f^{(s+\rho)} \left(- \frac{\partial x_{s+\rho}}{\partial x_\nu} \right) = \sum_{\rho=1}^{n-s} f^{(s+\rho)} D_{\rho\nu}, \quad \text{记 } D_{\rho\nu} = - \frac{\partial x_{s+\rho}}{\partial x_\nu}.$$

若对 $\nu > s$ 定义 $D_{\rho\nu} = \begin{cases} 1, & s+\rho=\nu \\ 0, & \text{否则} \end{cases}$, 则同样关系对 $\nu=s+1, \dots, n$ 也成立。因此

$$f^{(\nu)}(\mathbf{x}) = \sum_{\rho=1}^{n-s} T_\rho(f) D_{\rho\nu} \quad (\nu=1, \dots, n), \quad T_\rho(f) = f^{(s+\rho)}.$$

1 把上式移项：从 $f^{(\nu)} + \sum_{\rho} f^{(s+\rho)} \frac{\partial x_{s+\rho}}{\partial x_{\nu}} = 0$ 解出 $f^{(\nu)} = -\sum_{\rho} f^{(s+\rho)} \frac{\partial x_{s+\rho}}{\partial x_{\nu}}$ 。把负号收进新记号 $D_{\rho\nu} = -\partial x_{s+\rho} / \partial x_{\nu}$ ，就得到 $\nu \leq s$ 时的公式。

2 补全 $\nu > s$ 的情形：当 $\nu = s + \rho_0$ 是“被决定的坐标”时，它对应的偏导 $f^{(\nu)} = f^{(s+\rho_0)} = T_{\rho_0}(f)$ 。用刚定义的 $D_{\rho\nu}$ （只有 $\rho = \rho_0$ 那项为 1，其余为 0）正好让 $\sum_{\rho} T_{\rho}(f) D_{\rho\nu} = T_{\rho_0}(f) = f^{(\nu)}$ 。所以统一的公式对所有 $\nu = 1, \dots, n$ 都成立。

3 读出结构：对每一个在 $\mathcal{V}$ 上恒为零的 f ，它的整行梯度 $(f^{(1)}, \dots, f^{(n)})$ 都是同一组 $n - s$ 个向量 D_{ρ} 的线性组合，配方比例是 $T_{\rho}(f) = f^{(s+\rho)}$ 。关键： $D_{\rho\nu}$ 不依赖 f （只跟簇 $\mathcal{V}$ 的局部形状有关）， $T_{\rho}(f)$ 不依赖 ν 。

原文收尾：这就证明了命题中的结论，因为 $n - s \leq r - 1$ ，且这些关系在 $\mathcal{V}$ 上任一点成立，特别在 $\mathcal{X} \cap \mathcal{V}$ 的任一点成立。

1 每个 f_i （题设方程）在解集上、从而在 $\mathcal{V} \subseteq$ 解集 上处处为零，故可把上面的 f 取成 f_i ： $\frac{\partial f_i}{\partial x_{\nu}} = \sum_{\rho=1}^{n-s} T_{\rho}(f_i) D_{\rho\nu}$

2 记 $T_{i\rho} = T_{\rho}(f_i)$ 。由 4.2 的 $n - s \leq r - 1$ ，求和上限 $n - s$ 可以补零项延长到 $r - 1$ （多出来的 ρ 取 $T_{i\rho} = 0$ ）。于是

$$\frac{\partial f_i}{\partial x_{\nu}} = \sum_{\rho=1}^{r-1} T_{i\rho} D_{\rho\nu},$$

正是引理 14.2 的结论。它在 $\mathcal{V}$ 上任意一点成立，自然也在 $\mathcal{X}$ 落在 $\mathcal{V}$ 上的某点成立——这就找到了“某个点 $\mathbf{x}$ ”。■

第 5 节 引理 14.3：秩恰为 r 的 $\mathbf{x}$ 不超过 $R^{n-r+\varepsilon}$ 个

引理 14.3. 设对所有非零整点 $\mathbf{x} \neq \mathbf{0}$ 有 $C(\mathbf{x}) \neq 0$ 。则满足 $0 < |\mathbf{x}|$

证明的总策略（反证 + 引理 14.2） 假设这种 $\mathbf{x}$ 反而有 $\geq R^{n-r+\varepsilon}$ 个。这些 $\mathbf{x}$ 使一堆 $n - r + 1$ 阶子式为零（即矩阵 $M(\mathbf{x})$ 的秩 $= n - r$ ，由秩-零度定理等价于解空间维数 $= r$ ），正好落入引理 14.2 的框架。于是能找到一个特殊点 $\mathbf{x}$ ，在那里雅可比退化。把这点的退化信息，顺着第 2 节同

样的“求导—乘—求和”流程推下去，最终造出一个非零整点 $\mathbf{Y}$ 使 $C(\mathbf{Y}) = 0$ ，与“ C 不表示零”矛盾。所以那种 $\mathbf{x}$ 不可能那么多。

5.1 锁定研究对象 $\mathcal{X}$ ，套上引理 14.2

原文：所讨论的点是使矩阵 $(\sum_i c_{ijk} x_i)$ 秩恰为 r 的点。只需考虑整点集 $\mathcal{X}$ ：其中某个特定的 $n - r$ 阶子式 $\neq 0$ 、而所有 $n - r + 1$ 阶子式 $= 0$ ，并证 $|\mathcal{X}|$

关于“秩恰为 r ”的措辞（原文一处笔误） 原文这句写“矩阵 $(\sum_i c_{ijk} x_i)$ 的秩恰为 r ”，但紧接着的描述却是“某个 $n - r$ 阶子式非零、所有 $n - r + 1$ 阶子式全零”——后者意味着秩 $= n - r$ （而不是 r ）。这是 Davenport 原文的一处笔误：此处的秩应为 $n - r$ ，才与本引理“恰有 r 个线性无关解”（解空间维数 $= r$ ，由秩-零度定理得秩 $= n - r$ ）相吻合。核心是同一件事：把“矩阵秩 $= n - r$ ”这件几何条件，翻译成“某组 $n - r + 1$ 阶子式为零、某个 $n - r$ 阶子式不为零”这件多项式方程条件，从而能调用引理 14.2。下面跟随原文，记这组消失的 $n - r + 1$ 阶子式为 $f_1, \dots, f_N$ 。

1 为什么能“只看一个 $\mathcal{X}$ ”：“秩固定”可按“哪个 $n - r$ 阶子式非零”分成有限多类（子式的选法有限）。每一类是一个 $\mathcal{X}$ 。有限个集合里若总数 $\geq R^{n-r+\varepsilon}$ ，必有一类也几乎这么多，对那一类证矛盾即可。

2 套引理 14.2：在 $\mathcal{X}$ 上，所有 $n - r + 1$ 阶子式 $f_1, \dots, f_N$ 为零； N （子式个数）和它们的次数都被 n 限制住（子式是 $n - r + 1$ 个一次元的行列式，次数 $= n - r + 1 \leq n$ ）。假设 $|\mathcal{X}| \geq R^{n-r+\varepsilon}$ ，引理 14.2 给出一个特殊点 $\mathbf{x}^* \in \mathcal{X}$ ，在那里雅可比 $(\partial f_i / \partial x_\nu)$ 秩 $\leq r - 1$ 。这个 $\mathbf{x}^*$ 从此固定。

5.2 在每个 $\mathbf{x}$ 上造出 r 个线性无关解 $\mathbf{y}^{(p)}$

原文：对 $\mathcal{X}$ 中任意 $\mathbf{x}$ ，可像引理 14.1 那样，用某些 $n - r$ 阶子式当坐标，构造双线性方程的 r 个线性无关解 $\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(r)}$ 。（14.1 证明里只要一个解，取 $y_{n-r+1} = -\Delta$ 、其余为 0；推广到 r 个解是显然的。）

1 回忆 2.2: 当某个 $n - r$ 阶子式 $\Delta \neq 0$ 时, 把“多出来的那一列”换在不同位置, 可得到不同的解。秩 $= r$ 意味着零度 (解空间维数) $= r$, 所以确实能配出 r 个线性无关解。具体做法: 让额外的取值列在第 $n - r + 1, n - r + 2, \dots, n$ 这 r 个位置上轮流取 1, 各得一个解 $\mathbf{y}^{(p)}$, 它们的坐标都是 $\mathbf{x}$ 的 $n - r$ 阶子式 (整系数多项式)。

2 线性无关来自: 这些解在“被设为 1 的那一坐标”上互不相同地取了 1、其余取 0 (再乘 $-\Delta$), 所以谁也不是别人的组合。这一点在 5.4 末尾要紧。

5.3 建立恒等式, 并对它求导、配对

原文: 考虑任意点 $\mathbf{x}$ (实或复) 下 $B_j(\mathbf{x} | \mathbf{y}^{(p)})$ 的值, 它是开头那个矩阵的某个 $n - r + 1$ 阶子式 (有时会是两行相同的子式, 但无妨)。于是对任意 $\mathbf{x}$ 有恒等式

$$\sum_{i,k} c_{ijk} x_i y_k^{(p)} = \Delta_{j,p}(\mathbf{x}),$$

其中 $\Delta_{j,p}(\mathbf{x})$ 是某个 $n - r + 1$ 阶子式。当然, 若 $\mathbf{x} \in \mathcal{X}$, 所有 $\Delta_{j,p}$ 都为零。

1 左边即 $B_j(\mathbf{x} | \mathbf{y}^{(p)})$ 。把 $\mathbf{y}^{(p)}$ (其坐标是 $n - r$ 阶子式) 代进去, 行列式按列展开后整体是一个 $(n - r + 1) \times (n - r + 1)$ 的行列式, 即某个 $n - r + 1$ 阶子式 $\Delta_{j,p}(\mathbf{x})$ 。它对一切 $\mathbf{x}$ 都成立, 是多项式恒等式 (这一步关键: 不是只在 $\mathcal{X}$ 上, 而是处处成立, 才能放心求导)。

2 在 $\mathcal{X}$ 上 $\Delta_{j,p} = 0$: 因为 $\mathcal{X}$ 的定义就是“所有 $n - r + 1$ 阶子式为零”。

原文: 把 $x_1, \dots, x_n$ 当独立变量, 对 x_ν 求导:

$$\sum_k c_{\nu jk} y_k^{(p)} + \sum_{i,k} c_{ijk} x_i \frac{\partial y_k^{(p)}}{\partial x_\nu} = \Delta_{j,p}^{(\nu)}(\mathbf{x}),$$

右端 (ν) 表示偏导。乘以 $y_j^{(q)}$ ($1 \leq q \leq r$) 并对 j 求和, 得

$$\sum_{j,k} c_{\nu jk} y_k^{(p)} y_j^{(q)} + \sum_k \Delta_{k,q} \frac{\partial y_k^{(p)}}{\partial x_\nu} = \sum_j y_j^{(q)} \Delta_{j,p}^{(\nu)}(\mathbf{x}).$$

1 求导：和第 2 节 (*) 完全同型，只是右端不再是 0 而是 $\Delta_{j,p}^{(\nu)}$ （因为现在右端是子式 $\Delta_{j,p}$ ，不恒为零）。

2 乘 $y_j^{(q)}$ 求和，处理交叉项：第二个和里出现 $\sum_{i,j} c_{ijk} x_i y_j^{(q)}$ 。由 0.2 的换位公式它等于 $B_k(\mathbf{x} | \mathbf{y}^{(q)}) = \Delta_{k,q}(\mathbf{x})$ （用 5.3 第 1 步的恒等式，把上标换成 q ）。于是第二个和写成 $\sum_k \Delta_{k,q} \partial y_k^{(p)} / \partial x_\nu$ ，正是原文形式。这里不能像第 2 节那样直接说它为零，因为现在 $\mathbf{x}$ 还没固定、 $\Delta_{k,q}$ 一般非零；要等到把 $\mathbf{x}$ 取成 $\mathbf{x}^*$ 才让它为零。

5.4 对解空间里的一般点 $\mathbf{Y}$ 求和，固定 $\mathbf{x}^*$

原文：取由 $\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(r)}$ 张成的 r 维空间中任意点 $\mathbf{Y} = \sum_{p=1}^r K_p \mathbf{y}^{(p)}$ 。对它，用 $\sum_{p,q} K_p K_q (\dots)$ 把上式加权合并：

$$\sum_{j,k} c_{\nu jk} Y_j Y_k + \sum_{p,q} K_p K_q \sum_k \Delta_{k,q} \frac{\partial y_k^{(p)}}{\partial x_\nu} = \sum_{p,q} K_p K_q \sum_j y_j^{(q)} \Delta_{j,p}^{(\nu)} = \sum_p$$

对 $\nu = 1, \dots, n$ 成立。

符号卡 $\cdot \mathbf{Y} = \sum_p K_p \mathbf{y}^{(p)}$ 与 $A_{j,p}$

$\mathbf{Y}$ 是把 r 个基本解 $\mathbf{y}^{(p)}$ 以系数 $K_1, \dots, K_r$ 线性组合得到的任意解。它的分量 $Y_j = \sum_p K_p y_j^{(p)}$ 。约定 $A_{j,p} = \sum_q K_p K_q y_j^{(q)} = K_p \sum_q K_q y_j^{(q)} = K_p Y_j$ （把对 q 的和先收成 Y_j ）。这就是右端那步“记作如此”的来历。

1 左边第一项变 C -型：

$$\sum_{p,q} K_p K_q \sum_{j,k} c_{\nu jk} y_k^{(p)} y_j^{(q)} = \sum_{j,k} c_{\nu jk} \left(\sum_p K_p y_k^{(p)} \right) \left(\sum_q K_q y_j^{(q)} \right) = \sum_{j,k} c_{\nu jk} Y_j Y_k$$

。即把两组求和分别收成 Y_k, Y_j 。

2 右边收成 $A_{j,p}$ ：

$$\sum_{p,q} K_p K_q \sum_j y_j^{(q)} \Delta_{j,p}^{(\nu)} = \sum_{p,j} \left(\sum_q K_p K_q y_j^{(q)} \right) \Delta_{j,p}^{(\nu)} = \sum_{p,j} A_{j,p} \Delta_{j,p}^{(\nu)}$$

。

3 固定 $\mathbf{x} = \mathbf{x}^*$ （引理 14.2 找到的点）：这里 $\Delta_{k,q} = 0$ （所有 $n - r + 1$ 阶子式为零），所以左边第二项整块消失；同时由引理 14.2， $\Delta_{j,p}^{(\nu)} = \sum_{\rho=1}^{r-1} T_{j,p,\rho} D_{\rho\nu}$ （把 f_i 取成子式 $\Delta_{j,p}$ ）。代入得

$$\sum_{j,k} c_{\nu jk} Y_j Y_k = \sum_p \sum_j A_{j,p} \sum_{\rho=1}^{r-1} T_{j,p,\rho} D_{\rho\nu}.$$

5.5 乘 Y_ν 求和得 $C(\mathbf{Y})$, 再用 K 把它逼成 0

原文: 乘以 $Y_\nu = \sum_\sigma K_\sigma y_\nu^{(\sigma)}$ 并对 ν 求和:

$$C(\mathbf{Y}) = \sum_p \sum_j A_{j,p} \sum_{\sigma=1}^r \sum_{\rho=1}^{r-1} K_\sigma y_\nu^{(\sigma)} T_{j,p,\rho} D_{\rho\nu}.$$

(右端隐含对 ν 求和。) 现在选 $K_1, \dots, K_r$ 满足

$$\sum_{\sigma=1}^r K_\sigma \sum_\nu y_\nu^{(\sigma)} D_{\rho\nu} = 0 \quad (1 \leq \rho \leq r-1).$$

1 左边变 $C(\mathbf{Y})$:

$\sum_\nu Y_\nu \sum_{j,k} c_{\nu jk} Y_j Y_k = \sum_{\nu,j,k} c_{\nu jk} Y_\nu Y_j Y_k = C(\mathbf{Y})$ (三下标求和正是 C 的定义)。

2 右边整理: 把 $Y_\nu = \sum_\sigma K_\sigma y_\nu^{(\sigma)}$ 代入并对 ν 求和, 关键是把含 ν 的因子聚到一起: $\sum_\nu y_\nu^{(\sigma)} D_{\rho\nu}$ 。于是

$$C(\mathbf{Y}) = \sum_{p,j} A_{j,p} \sum_{\rho=1}^{r-1} T_{j,p,\rho} \left(\sum_{\sigma=1}^r K_\sigma \underbrace{\sum_\nu y_\nu^{(\sigma)} D_{\rho\nu}}_{\text{只含 } \sigma, \rho} \right).$$

3 让最内层括号为零: 方程组

$$\sum_{\sigma=1}^r K_\sigma \left(\sum_\nu y_\nu^{(\sigma)} D_{\rho\nu} \right) = 0, \quad \rho = 1, \dots, r-1$$

是关于 r 个未知数 $K_1, \dots, K_r$ 的 $r-1$ 个齐次线性方程。未知数比方程多一个, 必有非零解 (齐次方程组里方程数少于未知数时总有非零解)。

4 取成整数: 系数 $\sum_\nu y_\nu^{(\sigma)} D_{\rho\nu}$ 可取有理 ($D_{\rho\nu}$ 是局部坐标的导数, 可安排为有理数; $y_\nu^{(\sigma)}$ 是整数)。有理系数的齐次方程组, 把非零有理解通分即得非零整数解 $K_1, \dots, K_r$ 。

- 5 结论 $C(\mathbf{Y}) = 0$: 一旦每个 ρ 的内层括号为零, 第 2 步右端整体为 0, 故 $C(\mathbf{Y}) = 0$ 。

原文收尾: 于是 $\mathbf{Y} = \sum_p K_p \mathbf{y}^{(p)}$ 是非零整点 (因 $\mathbf{y}^{(p)}$ 线性无关、 K 不全为零), 又 $C(\mathbf{Y}) = 0$, 与 “ C 不表示零” 矛盾。结论得证。■

- 1 $K_1, \dots, K_r$ 为整数, $\mathbf{y}^{(p)}$ 坐标为整数, 故 $\mathbf{Y}$ 是整点。
- 2 $\mathbf{y}^{(p)}$ 线性无关 (5.2 末), 且 K 不全为零, 所以 $\mathbf{Y} \neq \mathbf{0}$ 。这一步用到了 “ $\mathbf{x}^* \in \mathcal{X}$ ” (在 $\mathcal{X}$ 上秩恰为预期值, r 个解才真线性无关)。
- 3 得到非零整点 $\mathbf{Y}$ 使 $C(\mathbf{Y}) = 0$, 推翻 “ C 不表示零”。所以最初的反设 “ $|\mathcal{X}| \geq R^{n-r+\varepsilon}$ ” 不成立, 即这种 $\mathbf{x}$ 少于 $R^{n-r+\varepsilon}$ 个。■

注 (不可忽视的逻辑陷阱)。上述证明中, $K_1, \dots, K_r$ 的选取没有用到数 $A_{j,p}$ 。否则推理就是谬误, 因为 $A_{j,p} = K_p Y_j$ 本身依赖 K ——若选 K 时又去迁就 $A_{j,p}$, 就成了 “用结果定原因” 的循环论证。我们只用 $\sum_\nu y_\nu^{(\sigma)} D_{\rho\nu}$ (与 $A_{j,p}$ 无关) 来定 K , 干净地避开了循环。

为什么这套又长又绕的推导是必须的 读到这里你可能会问: 第 2 节那么短, 为什么第 5 节这么长? 区别在于: 引理 14.1 处理的是 “ $H \equiv 0$ (全局退化)”, 可以直接用一个解、一次求导就拼出 $C(\mathbf{y}) = 0$; 而引理 14.3 处理的是 “在某一片 $\mathcal{X}$ 上退化、但不是全局恒等”, 此时子式 $\Delta_{j,p}$ 不恒为零, 求导后会冒出讨厌的右端项 $\Delta_{j,p}^{(\nu)}$ 。正是为了驯服这些非零的导数项, 才需要引理 14.2 提供 “雅可比秩 $\leq r-1 \implies$ 导数能用 $r-1$ 个 D 配出来”, 再用 $r-1$ 个方程定 K 把它们一举清零。长度来自这份额外的 “清理工作”。

第 6 节 引理 14.4: 备择情形 A 蕴涵 C 表示零 (收官)

引理 14.4. 引理 13.4 的备择情形 A 蕴涵 $C(\mathbf{x})$ 表示 0。

原文证明: 若 $C(\mathbf{x})$ 不表示零, 则引理 14.1 与 14.3 蕴涵: 存在 $\ll R^{n-r+\varepsilon}$ 个 $\mathbf{x}$ 使双线性方程关于 $\mathbf{y}$ 恰有 r 个线性无关解。因此存在 $\ll R^{n+\varepsilon}$ 对满足 \

$(0 < |\mathbf{x}|$

这是把前面所有引理“算总账”的一步。逐步看清那两个相乘的幂次怎么恰好凑成 $R^{n+\varepsilon}$ 。

1 反设：假设 C 不表示零。这样引理 14.1（保证 $H \neq 0$ ）与引理 14.3（计数）的前提都满足。

2 按秩分类数点对：每个满足双线性方程的整点对 $(\mathbf{x}, \mathbf{y})$ ，其 $\mathbf{x}$ 必使 $M(\mathbf{x})$ 退化（有非零 $\mathbf{y}$ 解）。按解空间维数 $r = 1, 2, \dots, n-1$ 把这些 $\mathbf{x}$ 分类。

3 固定 r ，数 $\mathbf{x}$ ：由引理 14.3，“恰有 r 个线性无关解”的 $\mathbf{x}$ 有 $\ll R^{n-r+\varepsilon}$ 个。

4 固定这样一个 $\mathbf{x}$ ，数 $\mathbf{y}$ ：满足 $M(\mathbf{x})\mathbf{y} = \mathbf{0}$ 的整数 $\mathbf{y}$ 构成一个秩 r 的格（解空间维数 $= r$ ）。在 $|\mathbf{y}|$

5 相乘：固定 r 时点对数 $\ll R^{n-r+\varepsilon} \cdot R^r = R^{n+\varepsilon}$ 。注意 R^{-r} 与 R^{+r} 正好抵消！——“ $\mathbf{x}$ 越退化越稀少（ R^{-r} ）”与“每个退化 $\mathbf{x}$ 配的 $\mathbf{y}$ 越多（ R^{+r} ）”精确平衡，这正是第 3 节那张图想说的事。

6 对 r 求和： r 只取 $1, \dots, n-1$ （有限项，项数 $\leq n$ ），每项 $\ll R^{n+\varepsilon}$ ，加起来仍 $\ll R^{n+\varepsilon}$ 。故满足双线性方程的整点对总数 $\ll R^{n+\varepsilon}$ 。

7 撞上情形 A 得矛盾：情形 A 断言这种点对“多于 $R^{n+\varepsilon_A}$ 个”。把本章用的 ε 取成 $\varepsilon = \frac{1}{2}\varepsilon_A$ ，则本章得到的上界 $R^{n+\frac{1}{2}\varepsilon_A}$ 当 R 充分大时小于情形 A 要求的下界 $R^{n+\varepsilon_A}$ 。一个量既要 $> R^{n+\varepsilon_A}$ 又要 $\ll R^{n+\varepsilon_A}$

8 结论：反设“ C 不表示零”不成立。故在情形 A 下， C 必表示零——存在非零整点 $\mathbf{z}$ 使 $C(\mathbf{z}) = 0$ 。■

“ $\ll R^n$ ”那句括注是什么意思 原文最后说“本证明实际表明点对数 $\ll R^n$ ”。看第 5、6 步：固定 r 时上界其实是 $R^{n-r+\varepsilon} \cdot R^r = R^{n+\varepsilon}$ ，这里的 ε 只来自引理 14.3 的计数余量。若改用本章开头提到的“更弱假设 AR^n ”那套写法，可把余量收紧到纯 R^n （不带 ε ）。这呼应了第 1 节的取舍说明： ε 只为行文方便，本质结果是 R^n 量级。



总点对数 $\ll \sum_r R^{n-r+\epsilon} \cdot R^r = R^{n+\epsilon} <$ 情形 A 的下界 $\Rightarrow$ 矛盾

结论：存在非零整点 z , $C(z)=0$

四条引理的依赖关系：14.1、14.2 是基石，14.3 借它们完成计数，14.4 把计数与情形 A 对撞出矛盾，逼出“C 表示零”。

回头看：本章在整个证明里的位置 第 13 章用圆法把“ $C(\mathbf{x}) = 0$ 有解吗”逼到引理 13.4 的岔路口：**情形 A**（双线性方程解多到反常）或**情形 B**（指数和有非平凡上界，可继续圆法）。本章证明的引理 14.4 表明：走情形 A 这条路，直接到达终点“C 表示零”。剩下的情形 B 由第 15–18 章用主弧/次弧、奇异积分、奇异级数、 p -adic 问题等圆法机器处理。两路合流，便得到 Davenport 关于**足够多变量的三次型必有非平凡零点**的著名定理。历史上，这类结果（Davenport 1959、1963 及后续 Davenport–Lewis 等）确立了“变量够多时三次方程一定有整数解”，是解析数论用圆法攻克**高次型**的里程碑；本章这套“用解的反常多迫出退化、再用代数几何收尾”的技巧，正是其中不依赖圆法、纯代数的关键一环。

返回 [全书目录](#)