數學傳播　44卷4期, pp. 78-86

答薛昭雄教授──並附評論與反思 (續)

林開亮

尊敬的薛教授:

謝謝您對拙文 [8] 的評論和建議 [9]。我在 [9] 中曾與您分享了我後續的一些想法。很高興又得到您進一步的回饋。這裡我想跟您分享一下近來我對 Heaviside 運算元法所得到的一些新的認識。在此之前, 我想就您的兩封信做個回饋。

在您的第一封信 [9, p. 91], 您對方程

$$250x \equiv 1 \pmod{2017} \tag{1}$$

的處理確實只用到了矩陣變換。蒙您指引, 這個方法更早見於 [13] 和 [11, pp. 203∼208]。其基礎是下述結果 ([8, p. 77] 練習 1, 證明平行於本文引理 1 的證明):

定理 1. 設二階整數矩陣

$$A = \begin{bmatrix} a & 1 \\ b & 0 \end{bmatrix}.$$

(i) 若二階整數矩陣 B 使得

$$BA = \begin{bmatrix} 1 & u \\ * & * \end{bmatrix},$$

則 $x = u$ 滿足同餘方程 $ax \equiv 1 \pmod{b}$。

(ii) 若二階整數矩陣 C 使得

$$CA = \begin{bmatrix} * & * \\ 1 & u \end{bmatrix},$$

則 $x = u$ 滿足同餘方程 $ax \equiv 1 \pmod{b}$.

這確實是一個矩陣方法：左乘相當於對矩陣做列變換 (elementary row operation)。但如何通過對矩陣 A 做變換? 這其實就需要用到帶餘除法：正是矩陣 A 的第一行的兩個元素 a, b 輾轉相除決定了對應的列變換。

比如, 在您所討論的方程 (1) 的情形, 我會把您的求解過程寫成這樣的形式 (其中箭頭 $\longrightarrow$ 下方的 r_1, r_2 分別表示箭頭前的那個矩陣的第一列與第二列) :

$$\begin{bmatrix} 250 & 1 \\ 2017 & 0 \end{bmatrix} \xrightarrow[r_2 \to r_2 - 8r_1]{2017 = 8\cdot 250 + 17} \begin{bmatrix} 250 & 1 \\ 17 & -8 \end{bmatrix} \xrightarrow[r_1 \to r_1 - 15r_2]{250 = 15\cdot 17 - 5} \begin{bmatrix} -5 & 121 \\ 17 & -8 \end{bmatrix} \xrightarrow[r_2 \to r_2 + 3r_1]{17 = (-3)\cdot(-5) + 2} \begin{bmatrix} -5 & 121 \\ 2 & 355 \end{bmatrix}$$
$$\xrightarrow[r_1 \to r_1 + 3r_2]{-5 = (-3)\cdot 2 + 1} \begin{bmatrix} 1 & 1186 \\ 2 & 355 \end{bmatrix}.$$

從而得到 $x = 1186$ 是同餘方程 (1) 的一個特解。請注意到, 每一步的列變換由前一個矩陣的第一行兩個元素的帶餘除法 (箭頭 $\longrightarrow$ 上方的算式) 所決定。只不過, 您這裡帶餘除法中的餘數取自除數 m 的一個絕對值最小的剩餘系 (從而除數和餘數都可以取負值)。

上述演算法中, 一旦出現餘數 ± 1, 演算法就終止, 此時它右方的整數除以這個絕對值最小的非零餘數, 就得到對應方程的一個特解; 否則 (出現的絕對值最小非零餘數不是 ± 1) 對應的同餘方程無解。誠如您所指出的, 這個方法可以用來確定整數 a, b 的最大公因數, 也可以用來求解丟番圖方程 $ax + by = c$。推而廣之, 這個方法可以用來確定一個整數矩陣 A 的不變因數, 也可以用 A 的 Smith 標準型求解丟番圖線性方程組 $AX = B$, 其中 X, B 是適當規格的整數向量。而這事實上是 MIT 數學教授 M. Artin 的經典代數教材 [1] 中的內容, 見該書 p.420 命題 14.4.9.

您在第二封來函中特別指出, 原始的 RSA 解密演算法中強調 $m = pq$ 是兩個不同素數的乘積, 這個確實很關鍵, 是我不應遺漏的。謝謝您的指正。我那裡主要是想說, 其中所涉及的模算術中的開 k 次方(參見 [12]) 與 [8] 中介紹的求微分/差分方程的特解之方法是平行的。

定理 2. 設 b, k, m 是給定的正整數, $\phi(m)$ 是 m 的歐拉函數, 滿足

$$\gcd(b,\ m) = 1 \qquad \text{且} \qquad \gcd(k,\ \phi(m)) = 1.$$

則下述步驟給出同餘方程

$$x^k \equiv b \pmod{m}$$

的解。

用求一術解關於 u 的同餘方程

$$ku \equiv 1 \pmod{\phi(m)},$$

得到 u, 然後令 $x \equiv b^u \pmod{m}$。

證明參見 [12]。事實上, Gauss 在他 1801 年出版的《數論研究》(見 [5]第 66 款) 中已經蘊含了這一結果。與 Gauss 的敘述平行的, 是下述結果:

定理 3. 設 A 是域 $\mathbb{F}$ 上的向量空間 V 上的綫性變換, $f \in V$。設存在 $\mu(x) \in \mathbb{F}[x]$ 使得 $\mu(A)f = 0$。給定 $P(x) \in \mathbb{F}[x]$, 若 $P(x), \mu(x)$ 互素, 則下述步驟給出方程

$$P(A)u = f$$

的解。

用求一術解關於 $U(x) \in \mathbb{F}[x]$ 的同餘方程

$$P(x)U(x) \equiv 1 \pmod{\mu(x)},$$

得到 $U(x)$, 然後令 $u = U(A)f$。

證明: 我們只要驗證 $u = U(A)f$ 滿足方程 $P(A)u = f$。由於 $P(x)U(x) \equiv 1 \pmod{\mu(x)}$, 可設 $P(x)U(x) = q(x)\mu(x) + 1$, 其中 $q(x) \in \mathbb{F}[x]$。從而

$$P(A)U(A) = q(A)\mu(A) + I,$$

其中 I 是 V 上的恒同變換。於是有

$$P(A)u = P(A)U(A)f = (q(A)\mu(A) + I)f = q(A)\mu(A)f + f = 0 + f = f. \qquad .$$

證畢。 □

注: 在 Jacobson 的綫性代數教程 [6, p. 67] 中, 零化 f 的非零多項式中次數最低的那個稱爲 f 的 order; 它相當於定理 2 中元素 b 模 m 的指數 (index)。定理 3 中的條件 $\mu(A)f = 0$, 相當於定理 2 中 b 滿足關係 $b^{\phi(m)} \equiv 1 \pmod{m}$, 這是著名的 Fermat–Euler 定理, 條件 $\gcd(b,\ m) = 1$ 就在這裡用到。

接下來, 我想著重跟薛老師與讀者分享一下我後來對 Heaviside 運算元法進一步思考所得到的一些認識。我總結成 *以下五點*。

第一: 指數平移定理 (Exponential Shift Theorem) 值得特別強調一下, 最好表述成以下形式:

定理 4 (指數平移定理)**.** 設 $P = P(x) \in \mathbb{C}[x]$, $\lambda \in \mathbb{C}$, 則對 $\mathbb{R}$ 中的任意可作用的函數 f, 有

$$P(D)\mathrm{e}^{\lambda x} f = \mathrm{e}^{\lambda x} P(D + \lambda) f. \tag{2}$$

證明: 我們先證明 (2) 對 $P(D) = D$ 成立。事實上, 根據 Leibniz 求導法則, 我們有

$$\begin{aligned} D(\mathrm{e}^{\lambda x} f) &= D(\mathrm{e}^{\lambda x})f + \mathrm{e}^{\lambda x} Df \\ &= \lambda \mathrm{e}^{\lambda x} f + \mathrm{e}^{\lambda x} Df \\ &= \mathrm{e}^{\lambda x}(D + \lambda) f. \end{aligned}$$

於是用數學歸納法, 我們可以證明, 對任意的正整數 n 有,

$$D^n(\mathrm{e}^{\lambda x}f) = \mathrm{e}^{\lambda x}(D+\lambda)^n f.$$

由此不難推出 (2) 對一切多項式 $P(D)$ 成立, 證畢。 □

它的第一個非平凡應用應該是推導下述結果:

定理 5. 設 $\lambda \in \mathbb{C}$, 則廣義特徵方程

$$(D-\lambda)^m u = 0 \tag{3}$$

的通解爲

$$u = \mathrm{e}^{\lambda x}(c_0 + c_1 x + \cdots + c_{m-1}x^{m-1}), \tag{4}$$

其中 $c_0, c_1, \ldots, c_{m-1}$ 爲任意常數。

證明: 根據指數平移公式 (2), 在變數替換 $u = \mathrm{e}^{\lambda x}v$ 之下, 方程 (3)等價於

$$D^m v = 0. \tag{5}$$

$m = 1$ 時, (3) 變成 $Dv = 0$, 其通解爲常值函數 $v = c$。當 $m > 1$ 時, 對 v 用帶餘項的 $m-1$ 階 Taylor 公式, 容易推出, 方程 (5) 的通解爲

$$v = c_0 + c_1 x + \cdots + c_{m-1}x^{m-1},$$

其中 $c_0, c_1, \ldots, c_{m-1}$ 爲任意常數。於是, 廣義特徵方程 (3) 的通解 $u = \mathrm{e}^{\lambda x}v$ 具有 (4) 的形式, 證畢。 □

第二: 可以將多項式的求一術及其矩陣變換基礎寫得更清楚, 如下。

引理 1. 設 $P(x), Q(x)$ 是非零多項式, 令

$$A = \begin{bmatrix} P(x) & 1 \\ Q(x) & 0 \end{bmatrix}.$$

(i) 若存在以多項式爲元素的二階矩陣 B 使得

$$BA = \begin{bmatrix} c & R(x) \\ * & * \end{bmatrix},$$

其中 c 爲非零常數, 則 $U(x) = \dfrac{R(x)}{c}$ 滿足同餘方程 $P(x)U(x) \equiv 1 \pmod{Q(x)}$。

(ii) 若存在以多項式爲元素的二階矩陣 C 使得

$$CA = \begin{bmatrix} * & * \\ c & R(x) \end{bmatrix},$$

其中 c 爲非零常數, 則 $U(x) = \dfrac{R(x)}{c}$ 滿足同餘方程 $P(x)U(x) \equiv 1 \pmod{Q(x)}$。

證明: 只證明 (i), (ii) 類似可證。設

$$B = \begin{bmatrix} S(x) & T(x) \\ * & * \end{bmatrix},$$

則計算可得

$$BA = \begin{bmatrix} S(x) & T(x) \\ * & * \end{bmatrix} \begin{bmatrix} P(x) & 1 \\ Q(x) & 0 \end{bmatrix} = \begin{bmatrix} P(x)S(x) + T(x)Q(x) & S(x) \\ * & * \end{bmatrix}.$$

於是, 根據假設, 我們有

$$\begin{bmatrix} P(x)S(x) + T(x)Q(x) & S(x) \\ * & * \end{bmatrix} = \begin{bmatrix} c & R(x) \\ * & * \end{bmatrix},$$

從而有

$$P(x)S(x) + T(x)Q(x) = c, \qquad S(x) = R(x).$$

將後一個式子代入前面的式子, 就有

$$P(x)R(x) + T(x)Q(x) = c.$$

因爲c爲非零常數, 等式兩邊除以c, 就得到

$$P(x)\frac{R(x)}{c} + \frac{T(x)}{c}Q(x) = 1.$$

這意味著 $U(x) = \dfrac{R(x)}{c}$ 滿足同餘方程 $P(x)U(x) \equiv 1 \pmod{Q(x)}$. (i) 證畢。 □

基於上述引理, 我們可以寫出求解同餘方程

$$P(x)U(x) \equiv 1 \pmod{Q(x)} \tag{6}$$

的秦九韶求一術如下(注意, 對矩陣 A 左乘一矩陣相當於對 A 做列變換, 初等矩陣對應著初等變換):

定理 6 (秦九韶求一術)**.** 給定同餘方程 (6), 其中 $P(x), Q(x)$ 是域 $\mathbb{F}$ 上的多項式, 且 $Q(x)$ 不是常數, $U(x)$ 是未知多項式。則可按以下步驟求出方程 (6) 的一個特解。首先寫出

$$A = \begin{bmatrix} P(x) & 1 \\ Q(x) & 0 \end{bmatrix}.$$

然後對第一行的兩個多項式用帶餘除法 (用次數高的多項式除以次數低的多項式), 設得到的商爲 $q(x)$, 則次數高的那一列減去次數低的那一列對應元素的 $q(x)$ 倍; 於是新得到的矩陣的第一行兩個元素替換爲第一次帶餘除法的除式與餘式, 重複之前的操作 (第一行兩個元素輾轉相除決定出各個列變換), 直到第一列某個數變成常數 c (演算法結束), 此時有下述結論:

(i) 方程 (6) 有解當且僅當 $c \neq 0$;

(ii) 當 $c \neq 0$ 時, 用 c 右邊的多項式除以 c, 就給出方程 (6) 的一個解 $U(x)$。

該演算法可以通過推廣 [8] 的證明得到, 並且從中可以看出: $c \neq 0$ 當且僅當 $P(x)$ 與 $Q(x)$ 互素, 此乃同餘方程 (6) 有解之充要條件。若 $c = 0$, 則與 c 同列的那個多項式是 $P(x)$ 和 $Q(x)$ 的最大公因式。

第三: 對 [8] 中所討論的常係數微分方程與差分方程, 可以列一個對照表(它可以視爲蔡聰明老師 [4] 中對照表的延續):

離散	連續
數列 $u(n)$	函數 $u(x)$
差分運算元 $\Delta : u(n) \mapsto u(n+1) - u(n)$	微分運算元 $D : u(x) \mapsto u'(x)$
$\Delta^{m+1}u = 0 \iff u$ 是 m 階以下等差數列 朱世傑招差公式(參見 [7, p. 67] 定理 3)	$D^{m+1}u = 0 \iff u$ 是 m 次以下多項式函數 Taylor 展開公式 (參見本文定理 5)
差分方程 $P(E)u = f$ 其中 $P(1) \neq 0$ 且 $\Delta^{m+1}f = 0$ 歸結爲 $P(E)U(E) \equiv 1 \pmod{(E-1)^{m+1}}$ 並令 $u = U(E)f$	微分方程 $P(D)u = f$ 其中 $P(0) \neq 0$ 且 $D^{m+1}f = 0$ 歸結爲 $P(D)U(D) \equiv 1 \pmod{D^{m+1}}$ 並令 $u = U(D)f$
$P(E)v(n) = \lambda^n f(n)$, 其中 $\Delta^{m+1}f = 0$ 令 $v(n) = \lambda^n u(n)$, 則根據 平移引理 $P(E)\lambda^n = \lambda^n P(\lambda E)$, 有 $P(\lambda E)u = f$	$P(D)v(x) = \mathrm{e}^{\lambda x} f(x)$, 其中 $D^{m+1}f = 0$ 令 $v(x) = \mathrm{e}^{\lambda x} u(x)$, 則根據 平移引理 $P(D)\mathrm{e}^{\lambda x} = \mathrm{e}^{\lambda x} P(D+\lambda)$, 有 $P(D+\lambda)u = f$

注: $E = \Delta + 1$ 是右平移運算元 (也可稱爲遞推算子) 在 [7, 8] 中用記號 T 表示, 但文獻中 E 更通用。

在歷史上, 微分方程與差分方程的理論一直平行發展, 例如 Heaviside 運算元法的先驅 George Boole (1815∼1864) 在 1859∼1860 年出版了《微分方程通論》[2] 和《差分方程通論》[3], 都曾作爲劍橋大學的教材。更讓人驚訝的是, Boole 在 [3, p.108] 中寫道:

> 只是在寫作本書時我才獲悉 Lobatto 先生的傑作《特徵理論》(*Théorie des Caractéristiques*), 它於 1837 年在阿姆斯特丹出版。該書包含了本節的內容, 以及微分方程中的類似定理, 一言以概之, 它囊括了其後一兩年在英格蘭重現發現的關於常係數線性微分方程的全部理論, 該理論發表在《劍橋數學雜誌》(*Cambridge Mathematical Journal*) 的前兩卷。每一個英國數學家將欣喜地看到我對 Lobatto 先生的公正。

這裡 Boole 提到的 Lobatto 先生是荷蘭數學家 Rehuel Lobatto (1797∼1866)。蒙好友吳帆老師告知：其著作《特徵理論》的完整名稱, 翻譯過來是《論數學分析中所用的特徵理論》, 其主題"特徵"即微分運算元與差分運算元。作者在書中提出, 可以將"特徵"與函數剝離開來獨立考慮, 可考慮"特徵"本身的多項式函數、甚至更一般的解析函數。

第四: 我們在 [8] 中對 Heaviside 工作的重新詮釋是站在秦九韶的角度考慮的。可以設想, Heaviside 本人未必認同我們用矩陣變換(定理6) 取而代之的做法。而且, 從一個教師的立場出發, 這個解法對沒有學過矩陣的學生來說是很難理解的。那麼, 有沒有一個更貼近 Heaviside 原始想法的處理呢? 實際上, 是有的。我們可以給出所需的多項式同餘方程 (見下面的 (7) 式, 它是 (6) 的特殊情況) 的一個特解公式, 如下:

定理 7. 設 $P(x)$ 是一個多項式, 常數項 $P(0) = 1$, 則同餘方程

$$P(x)U(x) \equiv 1 \pmod{x^{m+1}} \tag{7}$$

有特解

$$\boxed{U(x) = 1 + R(x) + R(x)^2 + R(x)^3 + \cdots + R(x)^m,} \tag{8}$$

其中

$$R(x) = 1 - P(x). \tag{9}$$

證明: 從 (9) 式得到 $P(x) = 1 - R(x)$, 從而就有

$$P(x)U(x) = [(1-R(x)][1+R(x)+R(x)^2+R(x)^3+\cdots+R(x)^m] = 1-R(x)^{m+1}. \tag{10}$$

由於 $R(x) = 1 - P(x) = P(0) - P(x)$, 因此 $R(x)$ 中每一個單項的次數都大於或等於 1, 從而 $R(x)^{m+1}$ 中的每個單項的次數都大於或等於 $m+1$, 這就表明 x^{m+1} 整除 $R(x)^{m+1}$ 中的每個單項, 從而 $x^{m+1} \mid R(x)^{m+1}$, 而 $-R(x)^{m+1} = P(x)U(x) - 1$, 這就意味著 $x^{m+1} \mid P(x)U(x) - 1$, 即有 (7) 式成立。證畢。 □

於是, 用定理 7 替代定理 6, 就可以解釋 Heaviside 演算法的有效性。這個觀點還有一個好處, 即它很容易推廣到求常係數偏微分方程的特解, 我們將在別處介紹。

第五: 我要跟薛教授分享的是, 我與中央民族大學王兢博士在合作論文 [10] 中, 對 Heaviside 運算元法, 曾給出一個基於形式冪級數等式

$$(1-x)(1+x+x^2+x^3+\cdots) = 1 \tag{11}$$

的理解, 並與 p-進數做了類比 (據說, Kurt Hensel 最初正是通過與無窮級數類比而發現了 p-進數)。而我在前面所分享的第四點則指出, 無窮級數完全可以避免 : 我們可以用 (11) 的有窮版本

$$(1-x)(1+x+x^2+x^3+\cdots+x^m) = 1 - x^{m+1} \tag{12}$$

來理解 Heaviside 運算元法 (這個觀察由丁玖教授向作者指出, 從而大大化簡了定理7的證明)。從某種意義上說, 這個理解是秦九韶矩陣方案跟 Heaviside 無窮級數方案的折中。慚愧的是, Heaviside 的原始工作我們尚未觸及, 是研究的不足。

薛教授, 感謝您對拙文關注, 啟迪我進一步思考, 從中享受許多樂趣。敬請先生多多指教。

順便說一句, 我後來注意到您與我極欽佩的數學家徐利治先生 (1920∼2019) 有過密切合作, 開亮有緣得您指點一二實屬三生有幸。疫情之下, 請您多多保重。

致謝

感謝美國南密西西比大學 丁玖 教授、中央民族大學王兢博士、學友吳帆老師與本文審稿老師對初稿提出寶貴意見。遵循審稿老師的意見, 作者對初稿做了刪減。

參考文獻

1. Michael Artin, *Algebra*, Pearson Education, Inc., 1991.
2. George Boole, *A Treatise on Differential Equations*, Cambridge, Macmillan, 1859.
3. George Boole, *A Treatise on the Calculus of Finite Differences*, Cambridge: Macmillan, 1860.
4. 蔡聰明。微積分與差和分大意 — 連續與離散之間的類推。數學傳播季刊, 2(2), 34-39, 1977。

5. C. F. Gauss, *Disquisitiones Arithmeticae*, translated by Arthur A. Clarke, Springer, 1986.
6. Nathan Jacobson, *Lectures in Abstract Algebra: II. Linear Algebra*, Graduate Texts in Mathematics 31, Springer-Verlag, New York, 1953.
7. 林開亮, 微積分之前奏 (或變奏): 高階等差數列的求和, 數學傳播季刊, 41(1), 61-79, 2017。
8. 林開亮, 解常係數線性微分方程和遞推關係的新方法 — 秦九韶和亥維賽的遺產。數學傳播季刊, 43 (2), 63-79, 2019。
9. 林開亮, 薛昭雄。薛昭雄來函暨林開亮回覆。數學傳播季刊, 43(4), 91-98, 2019。
10. 林開亮, 王兢。論 Heaviside 算子法的合理性。好玩的數學, 2020-02-20。
11. R. S. Millman, Peter J. Shiue and E. B. Kahn, *Problems and Proofs in Numbers and Algebra*, Springer International Publishing, Switzerland, 2015.
12. J. H. Silverman, *A Friendly Introduction to Number Theory*(4th Edition), Pearson Education, Inc. 2012. 孫智偉等 (譯)。《數論概論》。北京：機械工業出版社, 2016。
13. J. R. Silvester, A Matrix Method for Solving Linear Congruences, *Math. Mag.* 53(2), 90-92, 1980.

—本文作者任教中國西北農林科技大學理學院—