數學傳播　43卷4期, pp. 91-98

薛昭雄來函暨林開亮回覆

梁教授:

頃閱數學傳播第 170 期 (2019 年 6 月) 文章"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", 作者爲林開亮先生。這是一篇非常好的科普文章, 值得推介給有興趣的讀者。

然而, 在第五節: 解整數同餘方程的求一術中, 作者提到

"如解方程

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

你可能就無計可施了!"

本人覺得上面的敘述值得商榷, 因爲 (1) 之解可用矩陣方法 (不必利用整數之帶餘除法), 請見文獻 [1] 及 [2]。在 [2] 中 (請參閱 6.5 節) 已將 [1] 之方法推廣至求 $\gcd(a,b)$, a,b 爲正整數及求解線性不定方程 $ax+by=c$, 其中 a,b 與 c 均表非零整數且 $\gcd(a,b) \mid c$。

下面是(1) 式之解之說

$$\begin{bmatrix} 250 & 1 \\ 2017 & 0 \end{bmatrix} \xrightarrow[\text{加到第二列}]{\text{將第一列乘}(-8)} \begin{bmatrix} 250 & 1 \\ 17 & -8 \end{bmatrix} \xrightarrow[\text{加到第一列}]{\text{將第二列乘}(-15)} \begin{bmatrix} -5 & 121 \\ 17 & -8 \end{bmatrix}$$

$$\xrightarrow[\text{加到第二列}]{\text{將第一列乘}(+3)} \begin{bmatrix} -5 & 121 \\ 2 & 355 \end{bmatrix} \xrightarrow[\text{加到第一列}]{\text{將第二列乘}(+3)} \begin{bmatrix} 1 & 1186 \\ 2 & 355 \end{bmatrix}$$

即得 $x \equiv 1186 \pmod{2017}$。

參考文獻

1. John R. Silvester, A matrix method for solving linear congruences. *Math. Mag.* 53(2) (1980), 90-92.
2. Richard S. Millman, Peter J. Shiue and Eric Brendan Kahn, *Problems and Proofs in Numbers and Algebra*, Springer International Publishing Switzerland, 2015.

薛昭雄

美國內華達州立大學教授

答薛昭雄教授 — 並附評論與反思

尊敬的薛教授:

謝謝您對拙文 [1] 的評論和建議。特別感謝您指出的第 3 條意見, 因爲我那裡沒有表達清楚本意, 可能會讓讀者誤解。第5節的註腳5, 本意就是建議讀者用求一術解同餘方程 $250x \equiv 1 \pmod{2017}$, 正如您在來信中所指出的那樣。對於您提的其他三條意見, 我在關於中國古代數學史的通俗報告 [2] 中也曾提到, 但只是點到爲止, 未詳細展開。謝謝您告訴我相關文獻, 這是我之前沒有注意到的。

此外, 對於 [1], 我還有一些評論與反思, 想與您及其他讀者分享, 如下。請您批評指正!

[1] 探討微分方程

$$P(D)u = f$$

的求解, 其中 P 是一個複係數多項式, f 是一個擬多項式[1], u 是未知函數。也許應該指出, 只要 f 連續, 就可以通過累次積分求出 u。道理在於, 根據代數基本定理, $P(D)$ 可以分解爲線性因數的乘積 (不妨設 P 的首項係數等於 1):

$$P(D) = \prod_{i=1}^{n}(D - \lambda_i),$$

若對 $k = 1, \ldots, n$, 令 $u_k = (D - \lambda_k)u_{k-1}$, 而 $u_0 = u$, 則

$$u_n = \prod_{i=1}^{n}(D - \lambda_i)u_0 = P(D)u,$$

於是, 爲了求解方程 $P(D)u = f$, 我們只需對 $k = n, n-1, \ldots, 1$, 逐一求解

$$(D - \lambda_k)u_{k-1} = u_k,$$

其中 $u_n = f$。注意到, 每個方程都是一階線性微分方程, 都可直接求解。作爲例子, 我們來看 $k = n$ 的情況, 此時我們要求解的是

$$(D - \lambda_n)u_{n-1} = f,$$

根據一階線性微分方程的基本結果 (下面也有推導),

$$u_{n-1} = \mathrm{e}^{\lambda x}\int \mathrm{e}^{-\lambda x} f.$$

[1]形如 $\mathrm{e}^{\lambda x}f(x)$ (其中 $f(x)$ 爲多項式) 的函數, 稱爲擬多項式。當 $\lambda = 0$ 時, 我們得到真正的多項式 $f(x)$。

原則上據此可以依次求出 $u_{n-1}, \ldots, u_0 = u$。不過, 這個方法的缺陷在於, 必須確定 $P(D)$ 的因數分解; 這在通常情況下是很難辦到的。在 f 爲多項式或擬多項式的情形下, Euler 曾提出待定係數法來求解, 這是通常教科書裡會介紹的方法。Oliver Heaviside 則發展了運算元法, 其基礎則是無窮級數。[2]

[1]指出, 在 f 爲擬多項式的情形, 這本質上可歸結爲一個多項式同餘方程的求解, 而它可以通過直接應用秦九韶的求一術來實現。

其實這篇文章還有許多可以改進的地方, 這裡我特別想指出的有五點。

第一：[1] 建議的用求一術求解常係數線性微分方程或差分方程其實有一個平行的數論版本, 這就是著名的 RSA 解密算法(參見 [3, 4, 5]):

定理1：[RSA **解密算法**] 設 b, k, m 是給定的整數, m 是正整數, $\phi(m)$ 是其歐拉函數。

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

則下述步驟, 給出同餘式

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

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 $x = b^u \pmod{m}$。

第二：[1] 第 8 節的第 (15) 式 (也見第 11 節練習 2) 其實有個名稱, 叫 "指數平移定理" (Exponential Shift Theorem):

定理2：[**指數平移定理**] 設

$$P(D) = a_n(x)D^n + a_{n-1}(x)D^{n-1} + \cdots + a_1(x)D + a_0(x),$$

其中 $a_i(x)$ 是區間 I 上的函數, D 是對 x 的求導運算元。設 $\lambda(x)$ 是區間 I 上的光滑函數,

[2]另一種處理方式是 Laplace 變換, 但正如
Gian-Carlo Rota 在 Ten lessons I wish I had learned before I started teaching differential equations一文中指出的 (見 Lesson nine: Motivate the Laplace transform.)

> Ordinarily, we motivate the Laplace transform by appealing to initial value problems for linear differential equations with constant coefficients. But this motivation is rather thin: taking inverse Laplace transforms is no joke and initial value problems can be solved in other ways. I do not know how to properly motivate the Laplace transform ⋯

則在 I 上的 n 次可導函數空間上, 成立以下運算元等式:

$$\mathrm{e}^{-\lambda(x)}P(D)\mathrm{e}^{\lambda(x)} = P(D+\lambda'(x)),$$

其中

$$P(D+\lambda'(x)) = a_n(x)(D+\lambda'(x))^n + a_{n-1}(x)(D+\lambda'(x))^{n-1} + \cdots + a_1(x)(D+\lambda'(x)) + a_0(x).$$

特別地, 對 $\lambda(x) = \lambda x$ (此處 λ 爲常數), 我們有

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

注意, 如同正文中一樣, 此處 $\mathrm{e}^{-\lambda(x)}$ 與 $\mathrm{e}^{-\lambda(x)}$ 均表示由各自決定的左乘運算元, $P(D)$ 中的各個係數 $a_i(x)$ 也是如此理解。更值得注意的是, $(D+\lambda'(x))^k$ 是運算元 $D+\lambda'(x)$ 的 k 次冪, 因此要作爲運算元的複合來理解, 不能視爲普通的 (交換) 二項式展開。例如, 計算 $(D+x)^2$, 我們有

$$\begin{aligned}(D+x)^2 &= (D+x)(D+x)\\ &= D^2 + Dx + xD + x^2\\ &= D^2 + (xD+1) + xD + x^2\\ &= D^2 + 2xD + x^2 + 1,\end{aligned}$$

其中我們用到了著名的對易關係

$$Dx - xD = 1,$$

它只不過是求導的 Leibniz 法則的直接應用：對任意的可導函數 f 有

$$Dx(f) = D(x\cdot f) = D(x)\cdot f + x\cdot D(f) = f + xD(f) = (1+xD)f.$$

定理2的證明: 根據上述解釋, 我們只要證明對 $P(D) = D^k$, $k = 0, 1, \ldots, n$ 的情況證明即可, 事實上不難發現, 只要證明 $P(D) = D$ 的情況, 此時任取一個可導函數 f, 計算

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

這個定理的一個直接應用是, 求解原文第 9 節中的廣義特徵方程 (見 [1, p. 71] (17) 式) $(D-\lambda)^m y=0$, 容易看出, 在變數替換 $y=\mathrm{e}^{\lambda x}z$ 下, 上述方程等價於 $D^m z=0$, 後一個方程推出 z 的通解是次數小於 m 的多項式, 從而 y 具有形式

$$y=\mathrm{e}^{\lambda x}(C_{m-1}x^{m-1}+C_{m-2}x^{m-2}+\cdots+C_0),$$

即 y 爲次數小於 m 的擬多項式。

此外, 該定理還可以倒過來應用, 即我們利用

$$\mathrm{e}^{\lambda(x)}P(D+\lambda'(x))\mathrm{e}^{-\lambda(x)}=P(D)$$

來求解一個形如

$$P(D+\lambda'(x))y=f$$

的方程。容易看出在變數替換 $y=\mathrm{e}^{-\lambda(x)}z$ 下, 只需要求解方程

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

以下給出兩個例子:

例1: 求解 $(D+\lambda'(x))y=f$, 根據前面的說明, 令 $y=\mathrm{e}^{-\lambda(x)}z$, 只需要求解方程 $Dz=\mathrm{e}^{\lambda(x)}f$。積分得 $z=\int \mathrm{e}^{\lambda x}f$, 從而

$$y=\mathrm{e}^{-\lambda(x)}z=\mathrm{e}^{-\lambda(x)}\int \mathrm{e}^{\lambda(x)}f,$$

這就是我們在討論一階線性微分方程 $y'+py=q$ 時所得到的通解公式的等價表達。

例2: 前面我們已經算出 $(D+x)^2=D^2+2xD+x^2+1$, 於是可以討論方程

$$(D^2+2xD+x^2+1)y=f(x).$$

在變數替換 $y=\mathrm{e}^{-\frac{1}{2}x^2}z$ 下, 只需要求解方程

$$D^2z=\mathrm{e}^{\frac{1}{2}x^2}f.$$

連續積分兩次可得 z, 進而得出 $y=\mathrm{e}^{-\frac{1}{2}x^2}z$。

第三: [1] 建議的方法可以直接引出關於矩陣微分方程

$$X'=AX$$

的基本結果, 因爲在 Jordan 標準型下, 這個方程組 (經過變數替換 $X=PY$) 約化爲多個以下形式的獨立方程:

$$Y'=JY,$$

其中 J 是(由 A 確定的) 一個 m 階 Jordan 塊。當把這樣的方程明確寫出來之後, 不難發現 $Y=(y_1,\dots,y_m)^T$ 可以直接求解, 因爲 y_i 滿足 $y_{i+1}=(D-\lambda)y_i$, $i=1,\dots,m-1$, 從而 $y_m=(D-\lambda)^{m-1}y_1$, 而 $(D-\lambda)y_m=0$, 從而 $(D-\lambda)^{m-1}y_1=0$, 根據上面介紹的結果可以確定 y_1 的通解爲

$$y_1=\mathrm{e}^{\lambda x}\left(c_0+c_1x+c_2\frac{x^2}{2!}+\cdots+c_{m-1}\frac{x^{m-1}}{(m-1)!}\right),$$

其中 $c_0,\dots,c_{m-1}$ 爲任意複數。從而 Y 的通解爲

$$Y=\mathrm{e}^{\lambda x}\begin{pmatrix}1&x&\frac{x^2}{2!}&\cdots&\frac{x^{m-1}}{(m-1)!}\\0&1&x&\ddots&\frac{x^{m-2}}{(m-2)!}\\0&\ddots&\ddots&\ddots&\ddots\\\ddots&\ddots&\ddots&\ddots&x\\\ddots&\ddots&\ddots&\ddots&1\end{pmatrix}\begin{pmatrix}c_0\\c_1\\c_2\\\vdots\\c_{m-1}\end{pmatrix}=\Lambda(x)\cdot C,$$

其中 $C=(c_0,\dots,c_{m-1})^T\in\mathbb{C}^m$。注意, 矩陣 $\Lambda(x)$恰好是矩陣 xJ 的指數函數 $\Lambda(x)=\exp(xJ)$, 這裡

$$J=\begin{pmatrix}\lambda&1&0&\cdots&0\\0&\lambda&1&\ddots&0\\0&\ddots&\ddots&\ddots&0\\\vdots&\ddots&\ddots&\ddots&1\\0&0&0&0&\lambda\end{pmatrix}$$

是原始的 Jordan 塊。大多數常微分方程的教科書的處理過於複雜, 他們往往先定義以方陣 A 爲變數的指數函數 $\exp(A)$, 然後直接用公式 $X=\exp(At)C$ 給出方程 $X'=AX$ 的通解, 正如上面所指出的, 其實完全可以不提矩陣指數的分析概念。[3]事實上, [6] 就是先給出代數處理, 再介紹矩陣指數方法。不過, 這些作者並未指出, 可以用以上更簡單的方法來求 $Y'=JY$, 他們所用的仍然是待定係數法(見 [6, p.120])。

第四: 我們回顧一下上述求解思路, 會發現上述方法可用于求解形如

$$P(A)v=b$$

的方程, 其中 A 是某複綫性空間 V 上的線性運算元, b,v 分別是 V 上的已知向量與未知向量, P 是一個給定的複係數多項式。

[3]甚至可以由此來定義指數函數。

求解的基本假定是: 設 b 有零化多項式 $g(x)$。於是 b 落在空間 $\ker g(A)$, 設 $g(x) = (x-\lambda_1)^{m_1}\cdots(x-\lambda_s)^{m_s}$ 完全分解, 則根據正文 p.72 定理 3 (即 [5, 定理5.2.1]), $\ker g(A)$ 有直和分解:

$$\ker g(A) = \ker(A-\lambda_1)^{m_1} \oplus \cdots \oplus \ker(A-\lambda_s)^{m_s}$$

從而 $b = b_1 + \cdots + b_s$, 其中 $b_i \in \ker(A-\lambda_i)^{m_i}$, $i = 1, \ldots, s$。根據線性性質, 爲求解方程 $P(A)v = b$, 我們只需要對每個 b_i, 求解方程 $P(A)v = b_i$。換言之, 我們可以不妨假定 b 滿足 $(A-\lambda)^m b = 0$, 即 b 是 A 的屬於特徵值 λ 的廣義特徵向量。

於是, 我們考慮多項式同餘方程

$$P(x)U(x) \equiv 1 \pmod{(x-\lambda)^m}.$$

若 $P(\lambda) \neq 0$, 則上述方程有解, 並且可由求一術得到 $U(x)$; 從而令 $v = U(A)b$ 即可。

若 $P(\lambda) = 0$, 則不妨設 $P(x) = (x-\lambda)^n Q(x)$, 其中 $Q(\lambda) \neq 0$。我們可以分兩步來求解, 先解方程

$$(A-\lambda)^n u = b,$$

若它無解, 則原方程無解; 若它有解 u, 再用求一術求解 $Q(A)v = u$。注意向量 u 滿足

$$(A-\lambda)^{m+n}u = (A-\lambda)^m((A-\lambda)^n u) = (A-\lambda)^m b = 0,$$

而 $Q(\lambda) \neq 0$, 所以可以用求一術解出$Q(A)v = u$的解v, 並且它滿足 $(A-\lambda)^n(Q(A)v) = (A-\lambda)^n u = b$, 即 $P(A)v = b$。

綜上可見, 特解之所以能夠產生, 主要是因爲, 在最簡單的情況, b 是 A 的廣義特徵向量 (而在更一般的情況, b 是 A 的不同的廣義特徵向量之和)。例如, 當 $A = D$ 是微分運算元時, 非齊次項 b 是 D 的廣義特徵函數, 即擬多項式 (參見前文對定理2的第一個應用)。參見 [1, p.71] 的評注 1。

第五: 正如求一術可以推廣到求解同餘方程組, 這裡的方法也可以推廣到求解常係數線性微分方程組與常係數線性遞推關係組。我們只對於常微分方程組的情形, 說明大致思路。這一點已經爲 Bourbaki 寫進他們的實變函數著作, 參見 Bourbaki [8] 第 4 章第 2.9 節。

設我們考慮的方程組形如

$$AX = b,$$

其中 $A = (a_{ij})_{m\times n}$, 而每個元素 $a_{ij} = a_{ij}(D)$ 是關於 t 的求導運算元 $D = \frac{\mathrm{d}}{\mathrm{d}t}$ 的多項式, $b = (f_1, \ldots, f_m)^T$, 各個f_i 是 t 的已知向量值函數, $X = (x_1(t), \ldots, x_n(t))^T$ 是未知向量值函數。

求解思路如下, 設 D 爲未定元 λ, 將 A 視爲多項式環 $R=\mathbb{C}[\lambda]$ 上的矩陣, 根據主理想整環矩陣的一個基本結果 (Smith 標準型), 我們知道, 存在可逆矩陣 $P\in GL(m,R), Q\in GL(n,R)$ 使得

$$PAQ=\Lambda=\mathrm{diag}[d_1(\lambda),d_2(\lambda),\ldots,d_r(\lambda),0,\ldots,0],$$

其中 $d_1(\lambda),d_2(\lambda),\ldots,d_r(\lambda)$ 爲首一多項式, 且 $d_1(\lambda)\mid d_2(\lambda)\mid\cdots\mid d_r(\lambda)$。

在這個等式中將未定元 λ 替換爲 D, 得到

$$PAQ=\Lambda=\mathrm{diag}[d_1(D),d_2(D),\ldots,d_r(D),0,\ldots,0],$$

(其中 P,A,Q 我們用了同一個記號, 以避免符號複雜。)

注意到, 從 $AX=b$ 可以得到方程

$$PAX=Pb,$$

若我們令 $Y=Q^{-1}X$, 則它等價於

$$\Lambda Y=Pb.$$

注意到, 由於 Λ 是對角矩陣, 所以上述方程可以變數分離爲關於每個 y_i 的方程, 如果這樣的方程我們可以逐一判定它是否可解 (注意到若 Pb 是多項式或擬多項式, 則用 [1] 介紹的方法即可求解)。在最理想的情況, 我們可以求出各個 y_i, 從而定出 Y。現在我們要進一步求出 X。注意到, 從 $Y=Q^{-1}X$ 不難得到 $X=QY$, 現在 X 滿足 $PAX=Pb$, 由 P 可逆不難推出 X 滿足原來的方程 $AX=b$。

以上方法的一個數論版本可用於求解線性丟番圖方程組, 見[2]。

參考文獻

1. 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。
2. 林開亮。從《射雕英雄傳》到《九章算術注》。“好玩的數學” 微信公衆號。
3. 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。
4. 楊重駿、楊照昆。數論在密碼上的應用(下)。數學傳播季刊, 7(3), 2-7, 1983。在 數學知識網頁有全文的網頁版。
5. 林開亮。RSA 解密算法。數立方網站。
6. Morris W. Hirsch, Stephen Smale, Robert L. Devaney, *Differential Equations, Dynamical Systems, and an Introduction to Chaos*, Elsevier Academic Press, 2013.
7. 龔升, 張德健。線性代數五講　第五講: 向量空間在線性運算元下的分解。數學傳播季刊, 32(2), 34-53, 2008。
8. N. Bourbaki, *Functions of a Real Variable: Elementary Theory*, Springer, 2003.

—本文作者任教於中國西北農林科技大學—