感受伽羅瓦：排列與對稱多項式

在前面各章，筆者介紹了求二次至四次方程根式解的傳統方法，這些方法是人們經過數世紀累積經驗而總結出來的，每種方法都有其獨特性，除了都使用「契爾恩豪斯變換」(即設定 $x = y - \frac{b}{na}$，其中 n 是方程最高次項的冪，a 則是方程最高次項的係數) 外，便沒有其他共同點。但在 18 世紀，有數學家提出一種求二次至四次方程根式解的統一原理，此一原理須應用「排列」和「對稱多項式」的概念，因此在介紹此一原理前，須先介紹這兩個概念，這是本章的內容，而上述原理將留待下一章介紹。

首先介紹排列(permutation) 的概念。給定集合 X，X 的一個「排列」就是把 X 中元素排序的某個可能方案。舉例說，如果 $X = \{a, b, c\}$，那麼 acb 和 bca 就是 X 的兩個不同排列。在數學上，排列也可被看成 X 上的一個「一一到上函數」(one-one onto function)[1]。為了突出排列作為一種特殊函數的性質，我們可以把排列表達為兩行式(two line notation)，例如前述的兩個排列 acb 和 bca 便可以表示為

$$\begin{pmatrix} a & b & c \\ a & c & b \end{pmatrix} \qquad \begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix}$$

在上式中，第一行和第二行分別代表「定義域」和「對應域」(兩者皆為 X)，每一列的上下對應關係代表函數關係，例如左面的「兩行式」便代表把 a 映射為 a，把 b 映射為 c，把 c 映射為 b。根據組合數學，X 若包含 n 個元素，則 X 共有 $n!$ 個排列。

「兩行式」固然有直觀易明的優點，但在記寫上卻頗為費勁，因為每次都要寫出兩行內容重覆 (僅在次序上可能不同) 的元素。因此，數學家又創造出一種較簡便的記法，稱為循環式(cycle notation)。「循環式」的特點是僅用一行元素表達排列，該行元素由一或多個括弧組成。不同括弧中的元素各自獨立，彼此之間沒有映射關係；但在每個括弧中左右相鄰的任意

[1]設 f 為從「定義域」(domain) X 映射到「對應域」(codomain) Y 的函數，我們說 f 是「一一」(one-one) 的，當且僅當對 X 中任何兩個元素 x_1 和 x_2，如果 $x_1 \neq x_2$，則 $f(x_1) \neq f(x_2)$。f 是「到上」(onto) 的，當且僅當對 Y 中任何元素 y 而言，都有 X 中元素 x 使得 $f(x) = y$。

兩個元素之間則具有映射關係，即居於左邊的元素映射到居於右邊的元素。此外，還規定每個括弧中居於最右的元素映射到居於最左的元素。舉例說，前述的兩個排列 acb 和 bca 便可以表示為

$$(a)(b\,c) \qquad\qquad (a\,b\,c)$$

請注意在上面第一個「循環式」中，(a) 自成一個括弧，這代表把 a 映射為 a；$(b\,c)$ 則表示把 b 映射為 c，並把 c 映射為 b。上面第二個「循環式」則代表把 a 映射為 b，b 映射為 c，c 映射為 a。

對於「循環式」，有兩點須作說明。首先，由於各個括弧各自獨立而每個括弧實際上是圓圈，在寫出「循環式」時，各個括弧可以任意次序排列，而每個括弧可以隨意以任何一個元素作為開始。例如上面兩個「循環式」便可以分別改寫為

$$(c\,b)(a) \qquad\qquad (b\,c\,a)$$

其次，為盡量簡化寫法，我們約定「循環式」中任何只包含一個元素的括弧都可以略去不寫。這也就是說，我們約定任何沒有在「循環式」中寫出來的元素都各自組成一個括弧。舉例說，上面第一個「循環式」便可以改寫為

$$(c\,b)$$

當然，當我們想寫出恆等排列(identity permutation)(即把每個元素映射為自身的排列) 的「循環式」時，由於這個「循環式」的所有括弧都只包含一個元素，如果把所有括弧都略去不寫，我們的「循環式」將會是空的，所以我們又約定，對於「恆等排列」，它的「循環式」可表達為任意一個元素組成的括弧。舉例說，對於前述的 $X = \{a, b, c\}$ 來說，它的「恆等排列」便可以表達為以下三種「循環式」中的任何一種：

$$(a) \qquad\qquad (b) \qquad\qquad (c)$$

本網頁以下將主要使用「循環式」來表示排列。對於「排列」，本章的介紹到此為止，有關「排列」的其他內容，以後還會介紹。

接著介紹對稱多項式(symmetric polynomial) 的概念。前面各章討論的多項式都只包含一個變項，例如

$$x^2 + 2x + 1$$

中的 x，這種多項式稱為一元多項式(univariate polynomial)。但多項式其實可以包含多個變項，是為多元多項式(multivariate polynomial)，例如

$$x_1{}^3 + x_3{}^3 + 3x_1{}^2x_3 + 3x_1x_3{}^2 - 2x_2{}^2 + 5x_2 - 7 \qquad (1)$$

便包含 x_1、x_2 和 x_3 這三個變項。在一般情況下，給定一個 n 元多項式，可以把這個多項式的 n 個變項重新排列，從而得到一個新的多項式。舉例說，如果把 (1) 中的 x_2 和 x_3 對調，或者更形式化地說，對 (1) 進行 $(x_2\,x_3)$ 這個排列，便會得到以下這個新的多項式：

$$x_1{}^3 + x_2{}^3 + 3x_1{}^2x_2 + 3x_1x_2{}^2 - 2x_3{}^2 + 5x_3 - 7$$

但某些排列卻不會產生新的多項式，例如如對 (1) 進行 $(x_1\,x_3)$ 這個排列，會得到以下多項式：

$$x_3{}^3 + x_1{}^3 + 3x_3{}^2x_1 + 3x_3x_1{}^2 - 2x_2{}^2 + 5x_2 - 7$$

但上式實質上等同於 (1)。

如果某多項式在其變項經過任意排列後都是原來的多項式，那麼這個多項式稱為「對稱多項式」。容易看到，一元多項式總是對稱多項式，這是因為這種多項式只包含一個變項，而對一個變項只能進行一種排列－恆等排列。以下提供多元對稱多項式的例子：

$$9x_1{}^7x_2 + 9x_1{}^7x_3 + 9x_1x_2{}^7 + 9x_1x_3{}^7 - 4x_1 + 9x_2{}^7x_3 + 9x_2x_3{}^7 - 4x_2 - 4x_3 \qquad (2)$$

上式包含三個變項 x_1、x_2 和 x_3，對這三個變項的排列共有六種：(x_1)、$(x_1\,x_2)$、$(x_1\,x_3)$、$(x_2\,x_3)$、$(x_1\,x_2\,x_3)$、$(x_1\,x_3\,x_2)$。讀者可自行驗證，對 (2) 進行上述六種排列的任何一種，都會得到原來的多項式，所以 (2) 確是對稱多項式。

在眾多對稱多項式中，有一類稱為基本對稱多項式(elementary symmetric polynomial)，對方程根式解的研究有重要作用。設有 n 個變項 x_1、...、x_n，那麼由這些變項可構成 n 個基本對稱多項式 s_1、...、s_n，其定義如下[2]：

$$\begin{aligned}
s_1 &= \sum_{1\le n} x_j \\
&\vdots \\
s_k &= \sum_{1\le j_1<j_2<\cdots<j_k\le n} x_{j_1}x_{j_2}\cdots x_{j_k} \\
&\vdots \\
s_n &= x_1x_2\cdots x_n
\end{aligned}$$

[2]請注意 s_1、...、s_n 的定義其實依賴於所含的變項，例如當變項是 x_1、x_2 時，$s_1 = x_1 + x_2$；當變項是 y_1、y_2、y_3 時，$s_1 = y_1 + y_2 + y_3$，因此嚴格地說，基本對稱多項式的符號應包含變項符號，即寫成 $s_1(x_1, x_2)$、$s_1(y_1, y_2, y_3)$ 等。但為免使符號過於複雜，這裡採用簡單的寫法，讀者從上下文應能判斷當前基本對稱多項式的定義是建基於哪些變項。

上述定義頗為抽象，但用一個實例便應能讓讀者明瞭。設有四個變項 x_1、x_2、x_3 和 x_4，那麼我們有以下五個基本對稱多項式：

$$\begin{aligned} s_1 &= x_1 + x_2 + x_3 + x_4 \\ s_2 &= x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 \\ s_3 &= x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 \\ s_4 &= x_1x_2x_3x_4 \end{aligned}$$

從以上例子可以看到，s_k (其中 $1 \leq k \leq n$) 就是以下計算的結果：從 x_1、...、x_n 這 n 個變項中選 k 個出來，使被選中的 k 個變項的冪為 1，沒被選中的 $n-k$ 個變項的冪為 0，然後把這 n 個帶有上述規定的冪的變項相乘。根據組合數學，上述選取共有 $C(n,k) = \frac{n!}{k!(n-k)!}$ 種可能性，因此共可得到 $C(n,k)$ 個乘積，把這些乘積加起來，便得到 s_k。

基本對稱多項式可以說是對稱多項式的基本構件，這是因為有以下定理：

定理 1 (對稱多項式基本定理Fundamental Theorem of Symmetric Polynomials)：設 f 為包含變項 x_1、...、x_n 的對稱多項式，則 f 可以改寫成以 s_1、...、s_n 作為變項的多項式。

舉例說，

$$x_1{}^2 - 2x_1x_2 + x_2{}^2 \qquad (3)$$

是對稱多項式，容易看到，這個對稱多項式可以改寫成 $(x_1+x_2)^2 - 4x_1x_2$，因而可以用基本對稱多項式 s_1 $(= x_1 + x_2)$ 和 s_2 $(= x_1x_2)$ 改寫成 $s_1{}^2 - 4s_2$。上面的 (3) 是簡單的對稱多項式，所以容易把它改寫成以基本對稱多項式作為變項的多項式。

可是，當給定的對稱多項式結構複雜時，便不那麼容易改寫了。以下介紹一種通用的改寫方法，這種方法須依賴於多項式中各個項的一種「詞典序」(lexicographic order) 大小比較。設某多項式包含變項 x_1、...、x_n，而 $cx_1{}^{p_1} \cdots x_n{}^{p_n}$ 和 $dx_1{}^{q_1} \cdots x_n{}^{q_n}$ 是多項式中的其中兩項 (其中 c 和 d 是係數)，我們規定 $c_1x_1{}^{p_1} \cdots x_n{}^{p_n} > c_2x_1{}^{q_1} \cdots x_n{}^{q_n}$ 當且僅當 $p_k > q_k$，其中 k 是最小的數字 i 使得 $p_i \neq q_i$。換言之，在比較兩項的大小時，不用理會係數，先比較兩項中 x_1 的冪次，具有較高冪次的項便是較大的項；如果兩項中 x_1 的冪次相等，便比較兩項中 x_2 的冪次，如此類推。舉例說，根據上述原理，我們有 $-4x_1 > 9x_2{}^7x_3$ (請注意 $-4x_1 = -4x_1{}^1x_2{}^0x_3{}^0$，$9x_2{}^7x_3 = 9x_1{}^0x_2{}^7x_3{}^1$)。請讀者自行驗證，上面 (2) 的對稱多項式中的各項正是按上述「詞典序」從大到小排列的。

現在假設給定一個對稱多項式 P，首先找出 P 中 (按前述「詞典序」) 的最大項 $cx_1{}^{p_1}x_2{}^{p_2}x_3{}^{p_3}\cdots x_n{}^{p_n}$，然後設定 $g_1 = cs_1{}^{p_1-p_2}s_2{}^{p_2-p_3}s_3{}^{p_3-p_4}\cdots s_n{}^{p_n}$，其中 s_1、s_2、s_3、...、s_n 是基本對稱多項式。接著計算 $P-g_1$，$P-g_1$ 雖然可能會比 P 多了一些項，但 $P-g_1$ 中的最大項卻必然比 P 中的最大項小。接下來我們又如法泡製 $P-g_1$：先找出 $P-g_1$ 中的最大項，然後設定相應的 g_2，接著計算 $P-g_1-g_2$，如是者繼續進行下去，最終必有 $P-g_1-g_2-\cdots-g_k=0$。由此可得

$$P = g_1 + g_2 + \cdots + g_k$$

由於上式中的 g_1、g_2、...、g_k 都是以 s_1、...、s_n 作為變項的多項式，所以上式必然也是以 s_1、...、s_n 作為變項的多項式。

以下用兩個例子說明上述方法，首先考慮以下包含三個變項 y_1、y_2 和 y_3 的對稱多項式 (在下式中，ω_3 是 1 的主幅角為 $\frac{2\pi}{3}$ 的立方根，並且根據《感受伽羅瓦：二次方程與複數》中的「定理 2」，有 $1+\omega_3+\omega_3{}^2=0$)：

$$\begin{aligned} P_1 &= (y_1+\omega_3 y_2+\omega_3{}^2 y_3)^3+(y_1+\omega_3{}^2 y_2+\omega_3 y_3)^3 \\ &= 2y_1{}^3-3y_1{}^2y_2-3y_1{}^2y_3-3y_1y_2{}^2+12y_1y_2y_3-3y_1y_3{}^2+2y_2{}^3-3y_2{}^2y_3 \\ &\quad -3y_2y_3{}^2+2y_3{}^3 \end{aligned}$$

容易看到，上式中的最大項是 $2y_1{}^3$ (即 $2y_1{}^3y_2{}^0y_3{}^0$)，因此根據前述方法，設定 $g_{11}=2s_1{}^{3-0}s_2{}^{0-0}s_3{}^0=2s_1{}^3$。接著計算

$$\begin{aligned} P_1-g_{11} &= P_1-2s_1{}^3 \\ &= P_1-2(y_1+y_2+y_3)^3 \\ &= P_1-2y_1{}^3-6y_1{}^2y_2-6y_1{}^2y_3-6y_1y_2{}^2-12y_1y_2y_3-6y_1y_3{}^2 \\ &\quad -2y_2{}^3-6y_2{}^2y_3-6y_2y_3{}^2-2y_3{}^3 \\ &= -9y_1{}^2y_2-9y_1{}^2y_3-9y_1y_2{}^2-9y_1y_3{}^2-9y_2{}^2y_3-9y_2y_3{}^2 \end{aligned}$$

讀者可以看到，上式中的最大項 $-9y_1{}^2y_2$ (即 $-9y_1{}^2y_2{}^1y_3{}^0$) 的確比 P_1 中的最大項 $2y_1{}^3$ 小。根據前述方法，設定 $g_{12}=-9s_1{}^{2-1}s_2{}^{1-0}s_3{}^0=-9s_1s_2$。接著計算

$$\begin{aligned} P_1-g_{11}-g_{12} &= P_1-g_{11}-(-9s_1s_2) \\ &= P_1-g_{11}+9(y_1+y_2+y_3)(y_1y_2+y_1y_3+y_2y_3) \\ &= P_1-g_{11}+9y_1{}^2y_2+9y_1{}^2y_3+9y_1y_2{}^2+27y_1y_2y_3+9y_1y_3{}^2 \\ &\quad +9y_2{}^2y_3+9y_2y_3{}^2 \\ &= 27y_1y_2y_3 \end{aligned}$$

讀者可以看到，上式中的最大項 $27y_1y_2y_3$ (即 $27{y_1}^1{y_2}^1{y_3}^1$) 的確比 $P_1 - g_{11}$ 中的最大項 $-9{y_1}^2y_2$ 小。根據前述方法，設定 $g_{13} = 27{s_1}^{1-1}{s_2}^{1-1}{s_3}^1 = 27s_3$。接著計算

$$\begin{aligned} P_1 - g_{11} - g_{12} - g_{13} &= P_1 - g_{11} - g_{12} - 27s_3 \\ &= P_1 - g_{11} - g_{12} - 27y_1y_2y_3 \\ &= 0 \end{aligned}$$

至此求得

$$\begin{aligned} P_1 &= g_{11} + g_{12} + g_{13} \\ &= 2{s_1}^3 - 9s_1s_2 + 27s_3 \end{aligned}$$

接下來考慮以下包含四個變項 y_1、y_2、y_3 和 y_4 的對稱多項式：

$$\begin{aligned} P_2 &= (y_1+y_2)(y_3+y_4)(y_1+y_3)(y_2+y_4) \\ &\quad +(y_1+y_2)(y_3+y_4)(y_1+y_4)(y_2+y_3) \\ &\quad +(y_1+y_3)(y_2+y_4)(y_1+y_4)(y_2+y_3) \\ &= {y_1}^2{y_2}^2 + 3{y_1}^2y_2y_3 + 3{y_1}^2y_2y_4 + {y_1}^2{y_3}^2 + 3{y_1}^2y_3y_4 + {y_1}^2{y_4}^2 + 3y_1{y_2}^2y_3 \\ &\quad +3y_1{y_2}^2y_4 + 3y_1y_2{y_3}^2 + 6y_1y_2y_3y_4 + 3y_1y_2{y_4}^2 + 3y_1{y_3}^2y_4 + 3y_1y_3{y_4}^2 \\ &\quad +{y_2}^2{y_3}^2 + 3{y_2}^2y_3y_4 + 3y_2{y_3}^2y_4 + 3y_2y_3{y_4}^2 + {y_2}^2{y_4}^2 + {y_3}^2{y_4}^2 \end{aligned}$$

上式中的最大項是 ${y_1}^2{y_2}^2$ (即 ${y_1}^2{y_2}^2{y_3}^0{y_4}^0$)，因此根據前述方法，設定 $g_{21} = {s_1}^{2-2}{s_2}^{2-0}{s_3}^{0-0}{s_4}^0 = {s_2}^2$。接著計算

$$\begin{aligned} P_2 - g_{21} &= P_2 - {s_2}^2 \\ &= P_2 - (y_1y_2 + y_1y_3 + y_1y_4 + y_2y_3 + y_2y_4 + y_3y_4)^2 \\ &= P_2 - {y_1}^2{y_2}^2 - 2{y_1}^2y_2y_3 - 2{y_1}^2y_2y_4 - {y_1}^2{y_3}^2 - 2{y_1}^2y_3y_4 - {y_1}^2{y_4}^2 \\ &\quad -2y_1{y_2}^2y_3 - 2y_1{y_2}^2y_4 - 2y_1y_2{y_3}^2 - 6y_1y_2y_3y_4 - 2y_1y_2{y_4}^2 \\ &\quad -2y_1{y_3}^2y_4 - 2y_1y_3{y_4}^2 - {y_2}^2{y_3}^2 - 2{y_2}^2y_3y_4 - 2y_2{y_3}^2y_4 \\ &\quad -2y_2y_3{y_4}^2 - {y_2}^2{y_4}^2 - {y_3}^2{y_4}^2 \\ &= {y_1}^2y_2y_3 + {y_1}^2y_2y_4 + {y_1}^2y_3y_4 + y_1{y_2}^2y_3 + y_1{y_2}^2y_4 + y_1y_2{y_3}^2 \\ &\quad +y_1y_2{y_4}^2 + y_1{y_3}^2y_4 + y_1y_3{y_4}^2 + {y_2}^2y_3y_4 + y_2{y_3}^2y_4 + y_2y_3{y_4}^2 \end{aligned}$$

上式中的最大項是 ${y_1}^2y_2y_3$ (即 ${y_1}^2{y_2}^1{y_3}^1{y_4}^0$)，因此根據前述方法，設定 $g_{22} = {s_1}^{2-1}{s_2}^{1-1}{s_3}^{1-0}{s_4}^0 = s_1s_3$。接著計算

$$\begin{aligned} P_2 - g_{21} - g_{22} &= P_2 - g_{21} - s_1s_3 \\ &= P_2 - g_{21} - (y_1+y_2+y_3+y_4)(y_1y_2y_3 + y_1y_2y_4 + y_1y_3y_4 + y_2y_3y_4) \end{aligned}$$

$$
\begin{aligned}
&= P_2 - g_{21} - {y_1}^2 y_2 y_3 - {y_1}^2 y_2 y_4 - {y_1}^2 y_3 y_4 - y_1 {y_2}^2 y_3 - y_1 {y_2}^2 y_4 \\
&\quad - y_1 y_2 {y_3}^2 - 4 y_1 y_2 y_3 y_4 - y_1 y_2 {y_4}^2 - y_1 {y_3}^2 y_4 - y_1 y_3 {y_4}^2 \\
&\quad - {y_2}^2 y_3 y_4 - y_2 {y_3}^2 y_4 - y_2 y_3 {y_4}^2 \\
&= -4 y_1 y_2 y_3 y_4
\end{aligned}
$$

上式中的最大項是 $-4y_1y_2y_3y_4$ (即 $-4{y_1}^1{y_2}^1{y_3}^1{y_4}^1$)，因此根據前述方法，設定 $g_{23} = -4{s_1}^{1-1}{s_2}^{1-1}{s_3}^{1-1}{s_4}^1 = -4s_4$。接著計算

$$
\begin{aligned}
P_2 - g_{21} - g_{22} - g_{23} &= P_2 - g_{21} - g_{22} - (-4s_4) \\
&= P_2 - g_{21} - g_{22} + 4y_1y_2y_3y_4 \\
&= 0
\end{aligned}
$$

至此求得

$$
\begin{aligned}
P_2 &= g_{21} + g_{22} + g_{23} \\
&= {s_2}^2 + s_1 s_3 - 4s_4
\end{aligned}
$$

連結至數學專題
連結至周家發網頁