

線性代數珍寶十三則*

林開亮

介紹

當我做學生的時候,我們用的矩陣論的教材是 Bocher 的舊得可憐的書(我認為寫得一團糟),我在這個科目上花的大量時間當中,我的主要情緒是惱火達到憤怒...四、五年以後,在我已經取得博士學位,聽過了 von Neumann 講算子理論以後,我才真正開始懂得這個科目是講什麼的。

Halmos(哈爾莫斯, 1916–2006),《我要做數學家》第51–52頁

我相信, Halmos 學習線性代數的經歷不是獨一無二的,事實上,清華大學數學系的周堅教授在給研究生上微分流形這門課的時候說得更直白:

第一次學線性代數時我雖然考試成績不錯,但是聽老師說沒有人是第一次學線性代數就真正學懂的。我後來學習微分流形的時候才真正把線性代數搞明白了。

如果筆者沒有記錯的話,這是當時他講到微分流形理論中的 Morse 引理時所發表的感慨,因為當時需要用到二次型的慣性定理。

對此,筆者也深有同感,大學一年級學線性代數的時候似懂非懂的,只是後來在別的學科(如多重微積分、高等幾何、數理統計)裏應用矩陣時才慢慢體會到線性代數的實質與威力。

Halmos 在聽了 von Neumann 的講座以後不久就動手寫了他這一生最有影響的一本書 [8](準確地說,是 [8]的前身,見 [10, 第126–128頁])。根據他自己在 [10, 第550頁]的說法,《有限維向量空間》和《希爾伯特空間問題集》(即參考文獻中的 [8]和 [9])或許是他寫得最好的書。筆者正是從 [8]中第一次瞭解到譜定理的極端重要性(見 [8, 第155頁] 譜定理 (spectral

*這裏我們借用了 D. Carlson, C. R. Johnson, D. C. Lay 與 A. D. Porter 的一篇文章的名字, *Gems of exposition in elementary linear algebra*, The College Mathematics Journal, **23**:4(1992), 300–303. 中譯文《初等線性代數珍寶數則》,見《數學譯林》第12卷(1993年)第3期, 249–252. 這些作者還編輯了一本同名的書, *Linear Algebra Gems*, Math. Asso. Amer., Washington, DC, 2002.

theorem) 一節的引言), 並體會到線性代數的樂趣。這裏, 筆者將這幾年來學習線性代數的一點心得與讀者一起分享, 也算是作為對 Halmos [8] 的一個回報。筆者深信, Halmos 必定對 G. H. Hardy 的下述格言深深贊同¹:

美是首要的試金石: 醜陋的數學不可能永存。

筆者希望, 這裏所選擇的一些例子或多或少具有優美的特徵。

1. AB 與 BA 的特徵子空間

一般來說, 兩個矩陣的乘積是不可交換的, 但是有一個基本的事實說, AB 與 BA 最重要的數值不變量—特徵值是相同的。事實上, 我們有這樣的結果: 設 A, B 是同階矩陣, 則 AB 與 BA 具有相同的特徵多項式。特別的, 如果 λ 是 AB 的特徵值, 具有代數重數 k , 則 λ 也是 BA 的特徵值, 而且具有相同的代數重數 k 。問題本來可以就此打住, 但是在很多情形我們更感興趣的是特徵值的幾何重數而不是代數重數。因此, 我們或許可以像 Flanders [6] 一樣, 轉而考慮它們的幾何重數是否相同的問題。即, 對於 AB 的屬於特徵值 λ 的特徵子空間 $E_\lambda(AB) = \{\xi : AB\xi = \lambda\xi\}$ 以及 BA 的屬於特徵值 λ 的特徵子空間 $E_\lambda(BA) = \{\xi : BA\xi = \lambda\xi\}$, 我們要問其維數是否相等? 這就引導我們得到下面的結果。

定理1: 設 A, B 分別是 $m \times n$ 和 $n \times m$ 矩陣, $\lambda \neq 0$ 是 AB 的特徵值, 則 λ 也是 BA 的特徵值, 而且 $\dim E_\lambda(AB) = \dim E_\lambda(BA)$ 。

證明: 設 $\xi \in E_\lambda(AB)$, 即有 $AB\xi = \lambda\xi$, 兩邊用 B 同時作用就有 $BA(B\xi) = \lambda(B\xi)$, 即 $B\xi \in E_\lambda(BA)$, 即 $B(E_\lambda(AB)) \subset E_\lambda(BA)$ 。類似的, $A(E_\lambda(BA)) \subset E_\lambda(AB)$ 。於是, 複合算子 AB 是 $E_\lambda(AB)$ 到自身的一個線性變換, 容易看出這個限制下來的線性變換實際上是以 $\lambda \neq 0$ 為係數的伸縮變換。因此, 我們立即有 $AB(E_\lambda(AB)) = E_\lambda(AB)$, 注意到 $AB(E_\lambda(AB)) \subset AE_\lambda(BA) \subset E_\lambda(AB)$, 從而 $AE_\lambda(BA) = E_\lambda(AB)$, 由此得到 $\dim E_\lambda(BA) \geq \dim E_\lambda(AB)$ 。根據對稱性有, $\dim E_\lambda(AB) \geq \dim E_\lambda(BA)$ 。這就證明了 $\dim E_\lambda(AB) = \dim E_\lambda(BA)$ 。證畢。

註記: 通常對“ $\lambda \neq 0$ 是 AB 的特徵值, 則 λ 也是 BA 的特徵值”這一事實的證明是借助於熟知的行列式等式

$$\lambda^n |\lambda I_m - AB| = \lambda^m |\lambda I_n - BA|,$$

但是上述證明則完全是幾何的。或許, 這一證明很好地佐證了 Artin [2, 第14頁]的話²:

¹孫以豐譯, 見 M. A. Armstrong 《基礎拓撲學》(中譯本) 第1頁第一章引言, 人民郵電出版社, 北京, 2010年。

²趙春來譯, 見 N. L. Blggs 與 A. T. White 《置換群與組合結構》(中譯本) 第18頁第二章引言, 北京大學出版社, 1987年。

我的經驗是，一個用矩陣進行的證明，如果你拋開矩陣的話往往可以使這個證明縮短一半。有時，這一點是辦不到的，你需要計算一個行列式。

[2]這本書名字為《幾何代數》，言下之意就是說用幾何的方法、從幾何的角度來研究代數。這個觀點既是 Halmos [8]所倡導的觀點，也是本文強調的重點所在。

2. 保角變換

在歐氏空間中，如果 $x, y \in \mathbb{R}^n$ 不為零，則 x, y 之間的夾角可以用內積 $\langle \cdot, \cdot \rangle$ 定義為

$$\angle(x, y) = \cos^{-1} \frac{\langle x, y \rangle}{\sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}} \in [0, \pi],$$

根據 Cauchy-Schwarz 不等式，這個定義是有意義的。 \mathbb{R}^n 的可逆線性變換 T 稱為保角的，如果對任意的非零向量 $x, y \in \mathbb{R}^n$ 有 $\angle(Tx, Ty) = \angle(x, y)$ 。一個基本的事實是，“保角”就是“保形”（常常稱為“共形”），即保持形狀，其精確的數學含義就是相似。顯然，每個相似變換都是保角的。因此，我們提到的這個基本事實無非就是說其逆命題也成立，可以認為這是相似的基本定理（它不是別的，只是中學平面幾何中關於三角形相似的 AAA 定理的本質）：

定理 2: 歐幾裏得空間 \mathbb{R}^n 上的每一個保角變換 T 具有形式 $T = kS$ 其中 k 為非零常數， S 是 \mathbb{R}^n 的正交變換。

事實上，我們可以證明下述更強的結論：

定理 3: 設 T 是歐氏空間中的可逆線性變換，則以下三個條件等價：

- (i) T 是保角線性變換。
- (ii) T 保持任意兩個向量之間的垂直（正交）關係不變。
- (iii) T 是某個正交變換的非零常數倍。

這裏 (i) \Rightarrow (ii) 與 (iii) \Rightarrow (i) 是顯然的，而我們有兩種方法證明 (ii) \Rightarrow (iii)。第一種是借助於 Schmidt 正交化手續，將任意兩個向量變換為兩個垂直的向量，然後利用保持垂直關係做計算。第二種方法更富技巧性，是利用下述代數引理（見 [8, 第 123 頁] 練習 4 (d)）來正交化任意兩個向量。

引理 1: 在實內積空間中， $(x + y) \perp (x - y) \iff \|x\| = \|y\|$ 。

於是對任意的 $x, y \in V$ ，構造向量 $\|y\|x + \|x\|y$ 與 $\|y\|x - \|x\|y$ ，它們互相正交。然後可以利用假設展開證明，我們留給有興趣的讀者。

定理2與定理3其實是非常簡單而基本的事實，但是從一些文獻的處理來看，似乎這些作者並沒有說清楚這一點。尤其是 Spivak [25, 第4頁]習題8，作者似乎不知道保角變換可以用相似的概念來精確刻劃，而且中譯本譯者似乎也被作者弄糊塗了，他們在 [26, 第144-147頁]所附加的“習題解答與提示”中給出的證明也是不得要領。其他的作者則或者將這個事實推給線性代數(例如 [17, 第15-16頁])，或者將相似變換直接拿來作為保角變換的定義(如 [13, 第561-562頁] 附加習題 1)；總而言之，他們總是假定讀者在線性代數中了解了這個基本的事實。或許，這個結果有必要寫進線性代數的教科書。正如初中幾何的主題是全等和相似，保角變換應該得到與等距變換同等程度的強調，而且這兩個概念之間的關係應該得到澄清。

註：引理1表達了這樣一個熟知的幾何事實：平行四邊形為菱形當且僅當對角線互相垂直。注意到，其對偶 $x \perp y \iff \|x + y\| = \|x - y\|$ 表達的是一個與之對偶的幾何事實：平行四邊形為矩形當且僅當對角線等長。

3. 華羅庚引理的再次應用

在 [15]中筆者類比抽象代數中環同態的華羅庚引理提出了所謂的“線性代數中的華羅庚引理”(見 [15, 第41頁] 引理1)：

引理2: 設 B 是域 F 上向量空間 V 上的雙線性型，使得對任意的 $x, y \in V$ 或者有 $B(x, y) = B(y, x)$ 或者有 $B(x, y) = -B(y, x)$ ，則 B 是對稱的或反對稱的。

筆者進一步利用這個引理對線性代數中的下述基本結果(例如，見 [2, 第110-112頁] 或者 [30, 第385頁]) 給出了一個更簡單的概念性的證明。

定理4: 設 B 是域 F 上向量空間 V 上的雙線性型，使得由 $B(x, y) = 0$ 定義的正交關係是對稱的，即 $B(x, y) = 0 \iff B(y, x) = 0$ ，則 B 是對稱的或者反對稱的。

事實上，筆者後來注意到，應用同樣的想法，利用數學歸納法，我們可以將上述結果從雙線性映射推廣到多線性映射。也就是說，事實上我們可以證明下述結果：

定理5: 設 A 是域 F 上向量空間 V 上的 n 重線性函數，使得如果 $A(x_1, \dots, x_n) = 0$ ，則對任意的置換 $\sigma \in S_n$ ，有 $A(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = 0$ ，則 A 是對稱的或者反對稱的。

事實上，只有 $n = 3$ 的情況需要單獨證明，其他情況直接應用歸納假設很容易證明。對於 $n = 3$ 的情況我們有所謂的辮子引理(見 [3, 第一卷第224頁])：

引理3: 設 V 和 W 是特徵不等於 2 的域上的向量空間, 且 $A : V \times V \times V \rightarrow W$ 是一個三線性映射, 對於前兩個變量是對稱的, 對於後兩個變量是反對稱的, 則 $A = 0$ 。

注: 事實上, 隱藏在這個形式計算背後的基礎是關於三個元素的全置換群 S_3 的生成元 $\sigma_1 = (12)$, $\sigma_2 = (23)$ 之間的基本關係 $(\sigma_1\sigma_2)^3 = e$ 。³ 有興趣的讀者可以參考 [3, 第一卷第 30 頁]練習 15。

4. 冪等矩陣

從幾何上講, 最簡單的線性變換是投影變換, 它的代數表現就是冪等矩陣 ($P^2 = P$)。從代數上講, 冪等矩陣只有唯一的不變量—秩: 兩個冪等矩陣相似當且僅當它們的秩相等。關於這類最簡單的矩陣, 我們也可以說出一些不平凡的事實, 以下就是一個例子。

定理6: 設 P, Q 是有限維向量空間 V 上的投影變換, 且 $P - Q$ 與 $Q - P$ 僅有平凡的不動點, 則 P, Q 相似。

證明: 令 $A = Q - P - I, B = P - Q - I$, 則

$$PB = P(P - Q - I) = P^2 - PQ - P = -PQ,$$

$$AQ = (Q - P - I)Q = Q^2 - PQ - Q = -PQ,$$

於是 $PB = AQ$, 由條件可知 A, B 可逆, 從而 $Q = A^{-1}PB$, 即 Q 與 P 等價, 從而 P, Q 有相同的秩。由於秩是投影變換在相似下的唯一不變量, 所以 P, Q 相似。證畢。

當 V 為內積空間且 P, Q 為正交投影變換時, 由於秩也是正交投影變換在正交相似下的唯一不變量, 所以 P, Q 正交相似。上述定理是王進賢與吳培元 [28, 第 260 頁]定理 1.3 的一個推論。該定理有一個無限維版本, 可見於 Riesz 與 Sz.-Nagy [22, 第 268 頁]。

定理7: 設 P, Q 分別是 Hilbert 空間 H 到子空間 V 和 W 上的正交投影, 而且

$$\|P - Q\| < 1$$

則 V 與 W 等距, 即存在一個變換 $T : V \rightarrow W$, 它是保距的。

³這個關係有時候又表為 $\sigma_1\sigma_2\sigma_1 = \sigma_2\sigma_1\sigma_2$, 即所謂的 Artin 辮子群關係。

5. Cochran 定理的代數本質

線性代數中有許多結果都有著代數的本質，概率統計中的 Cochran 定理就是一個著名的例子。1960年代，Chipman 和 Rao⁴ 指出，1930年代發現的 Cochran 定理實則是一個簡單的矩陣定理，而它所討論的只不過是冪等矩陣的代數性質。他們將這個定理表述如下：

定理 8: 設 A_1, \dots, A_k 為域 F 上 n 階矩陣使得 $\sum_i A_i = A$ 是冪等的，且 $\sum_i \text{rank}(A_i) = \text{rank}(A)$ ，則 A_1, \dots, A_k 都是冪等矩陣，而且當 $i \neq j$ 時， $A_i A_j = 0$ 。

這個定理有許多證明，一個巧妙證明可見 [8, p. 109] 習題 5。或許 Waterhouse [29] 提供的下述證明是最簡單的一個，因為他發現這個線性代數結果其實有著環論的背景。他的證明由以下兩個引理構成。

引理 4: 記 $\mathcal{M} = M(n, F)$ 為 F 上 n 階矩陣的集合。設 $A_1, \dots, A_k \in \mathcal{M}$ 使得 $\sum_i \text{rank}(A_i) = \text{rank}(\sum_i A_i)$ ，則

$$\left(\sum_i A_i\right)\mathcal{M} = \bigoplus_i A_i\mathcal{M}.$$

證明: 令 $A = \sum_i A_i$ 。很明顯， $AM \subset A_1\mathcal{M} + \dots + A_k\mathcal{M}$ 。注意到對任意的 n 階矩陣 C ，集合 CM 中的矩陣的各列恰為 C 的各列的線性組合，因此 $\dim CM = n \cdot \text{rank}(C)$ ，從而有 $\dim AM = n \cdot \text{rank}(A)$ 以及 $\dim A_i\mathcal{M} = n \cdot \text{rank}(A_i)$ 。因此，條件 $\sum_i \text{rank}(A_i) = \text{rank}(A)$ 意味著 $\dim AM = \sum_i \dim A_i\mathcal{M}$ ，從而有 $AM = \bigoplus_i A_i\mathcal{M}$ 。證畢。

引理 5: 設 R 是一個帶單位元的環。 e 是 R 中的冪等元， $e = e_1 + \dots + e_k$ ，並且滿足 $eR = \bigoplus_{i=1}^k e_i R$ ，則 e_1, \dots, e_k 都是冪等元，而且當 $i \neq j$ 時， $e_i e_j = 0$ 。

證明: 我們有 $e_i = e_i \cdot 1 \in e_i R \subset eR$ ，從而不妨設 $e_i = e s_i$ 。於是 $e_i = e s_i = e^2 s_i = e(e s_i) = e e_i = (e_1 + \dots + e_k) e_i = \sum_j e_j e_i$ ，由於 $\bigoplus_{i=1}^k e_i R$ 是直和分解，所以 $e_i = e_i \cdot 1 = e_i e_i = e_i^2$ ，且對於 $j \neq i$ 有 $e_j e_i = 0$ 。證畢。

現在只要將引理 5 應用於環 $R = \mathcal{M} = M(n, F)$ ，並注意到引理 4 的結論使得引理 5 的條件成立，即可推出定理 8。或許我們可以借用 J. J. Sylvester 的一句話來評論 Waterhouse 的這個證明：

一種奇怪的智力現象，可以歸結為，證明普遍論點通常比對其中的個別情況作出證明

⁴見 J. S. Chipman and M. M. Rao, Projections, Generalized Inverses, and Quadratic Forms. *J. Math. Anal. Appl.* **9**(1964), 1-11.

簡單得多。

註：目前大陸最爲全面的線性代數教材 [23, 第 259–261 頁] 以例子的形式給出了這個定理，但是那裏給出的證明過於複雜。筆者在 2006 年曾寫信告訴 [23] 的作者（當時 [23] 的前身收入了這個結果作爲練習）這個問題的上述簡單證法，但沒有收到回覆，不過在那一年的北京大學數學院研究生招生考試的高等代數試卷中出現了這道題。最近，筆者再次與該書的作者丘維聲先生討論了這個問題，他認爲，上述證明固然簡潔，但是需要環論的抽象背景，這對於初學線性代數的大一學生來說或許有些困難。不過對於那些學習過抽象代數的學生來說，這個證明倒可以算作一個非常好的應用舉例。

6. von Neumann 特徵值問題

線性代數中最重要的結果當屬譜定理，特別是實對稱矩陣、Hermite 矩陣和酉矩陣的譜定理。具體情況下求出矩陣的譜分解就是要求解矩陣的特徵值與特徵向量，毫無疑問，這是線性代數中最核心的問題。然而，一般來說，求解特徵值問題並不是一件容易的事。von Neumann [19] 解過一個有意思的特徵值問題（這個問題也收入了 L. Lovász [18, 第 20 頁]，習題 1.29），我們介紹如下，用以挑戰有興趣的讀者。

問題：求實二次型

$$\sum_{\mu=1}^{n-1} (x_{\mu+1} - x_{\mu})^2 = x_1^2 + 2 \sum_{\mu=2}^{n-1} x_{\mu}^2 + x_n^2 - 2 \sum_{\mu=1}^{n-1} x_{\mu} x_{\mu+1}$$

所對應的對稱矩陣的特徵值。

註：每一個 n 元二次型 $f(x) = \sum_{i,j=1}^n a_{ij} x_i x_j$ 對應著唯一的 n 階對稱矩陣 $A = (a_{ij})$ ，使得 f 可以寫爲 $f(x) = x'Ax$ ，其中 $x = (x_1, \dots, x_n)'$ 。

答案： $\lambda_{\mu} = 2 - 2 \cos \frac{\mu\pi}{n} = 4 \sin^2 \frac{\mu\pi}{2n}$ ， $\mu = 0, 1, \dots, n-1$ 。

評註：正如張堯庭在 [31, 第 352–353 頁] 所回憶的，中國近代數學家許寶騫 (1910–1970) 曾經“也考慮過這個問題，但是沒有解決，因爲一個特定矩陣的特徵值很難求，而 von Neumann 卻把特徵值和特徵向量放在一起來求，這樣問題反而解決了，所以他從這裏學到了東西。”恰如 J. J. Sylvester 所說的⁵：The more to do or to prove, the easier the doing or the proof.

⁵引自 M. Artin, Algebra, Pearson Education, Inc., 1991, p.197.

7. Schur 特徵值問題

矩陣論的另一個有趣的應用是用於計算數論中的 Gauss 和

$$\sum_{s=0}^{n-1} \varepsilon^{s^2}, \quad (\text{其中 } \varepsilon = e^{i\frac{2\pi}{n}} \text{ 是 } n \text{ 次單位根}).$$

Gauss 首先確定了這一求和, 然而他的方法非常巧妙。⁶後來, I. Schur 發現了一種較為簡單的方法, 他注意到, 這個 Gauss 和可以表達為一個 n 階矩陣

$$S = (s_{jk}) = (\varepsilon^{(j-1)(k-1)}), \quad (j, k = 1, \dots, n)$$

的跡。Schur 求出了這個矩陣的所有特徵值, 從而確定了 Gauss 和。因為 Schur 的這一貢獻, 這一矩陣也被命名為 Schur 矩陣, 而且通常記為 S 。關於 Schur 特徵值問題的具體求法, 我們留給有興趣的讀者, 也可以參考 [14, 第 207–212 頁]。這裏我們只給出最後結果如下: S 只有四個不同的特徵值, $\sqrt{n}, -\sqrt{n}, i\sqrt{n}, -i\sqrt{n}$, 按照 n 模 4 取值的情況, 其重數分別為:

- (1) 若 $n = 4k + 1$, 則對應的重數分別為 $k + 1, k, k, k$ 。
- (2) 若 $n = 4k + 2$, 則對應的重數分別為 $k + 1, k + 1, k, k$ 。
- (3) 若 $n = 4k + 3$, 則對應的重數分別為 $k + 1, k + 1, k + 1, k$ 。
- (4) 若 $n = 4k + 4$, 則對應的重數分別為 $k + 2, k + 1, k + 1, k$ 。

於是我們最後可以得到 Gauss 的著名公式 (見 [14, 第 191 頁]):

定理 9: 設 n 是一個正整數, 則

$$\sum_{s=0}^{n-1} e^{i\frac{2\pi}{n}s^2} = \begin{cases} (1+i)\sqrt{n} & \text{當 } n \equiv 0 \pmod{4} \\ \sqrt{n} & \text{當 } n \equiv 1 \pmod{4} \\ 0 & \text{當 } n \equiv 2 \pmod{4} \\ i\sqrt{n} & \text{當 } n \equiv 3 \pmod{4} \end{cases}$$

定理 9 在歷史上具有極高的重要性, 它是 Gauss 在 1801 年 5 月記錄下的一則日記 (見 [4, 第 481 頁]), Gauss 利用這個結果給出了二次互反律的第五個證明 (他一生總共給出了八個證明)。但是, Gauss 為了確定出 Gauss 和的符號, 竟然費了好幾年時間, 他在 1805 年 8 月 30 日記錄下這樣的日記 (見 [4, 第 481 頁至 482 頁]): “經過四年多的不懈努力之後, 1801 年 5 月所記錄的富有魅力的定理的證明終於大功告成了。” 即便如此, Gauss 最後得到這個結果也是靈感

⁶例如, 可見 T. Nagell, *Introduction to Number Theory*, AMS Chelsea Publishing, 1964, pp.177-180.

的傑作。他在1805年9月3日寫給 Olbers 的一封信中如是說道：“最後，只是幾天以前，終於成功了（我想說，不是由於我苦苦的探索，而只是由於上帝的恩惠）。就像是閃電轟擊的那一剎那，這個謎解開了；我以前的知識，我最後一次嘗試的方法、以及成功的原因，這三者究竟是如如何聯繫起來的，我自己也未能理出頭緒來。”⁷

註： $U = \frac{1}{\sqrt{n}}S$ 是一個酉矩陣，它與離散 Fourier 變換有密切的關係，對此有興趣的讀者請參考 Taussky [27] 中關於 Schur 矩陣一節的論述。

8. von Neumann 恒等式

von Neumann 的主要貢獻可能是無限維空間的算子代數，但是他對矩陣論也作出了巨大的貢獻。前面我們已經提到過，他巧妙地求解了一個特徵值問題，他1940年代發現的小結果（見 [20]）也可以算得上一顆小珍珠（見 [8, 第134頁] 習題 4(d) 的提示）。

定理10: 對於同階複方陣 A 和 B , 有

$$\|AB^* - B^*A\|^2 - \|AB - BA\|^2 = \text{tr}[(A^*A - AA^*)(B^*B - BB^*)]$$

其中 $\|A\|^2$ 的定義為 $\|A\|^2 = \text{tr}(AA^*)$ 。

並由此得到結論（見 [8, 第134頁] 習題 4(d), 一個基於譜定理的概念性證明可見 [8, 第171頁] 定理 2):

定理11: 若 A, B 之一為正規矩陣⁸, 則 A, B 可換蘊含 A, B^* 可換。

註: 定理11在無限維也成立, 稱為 Fuglede 可換性定理, 見 [7, 第68頁]。而且, 事實上我們有更為一般的 Putnam-Fuglede 定理, 見 [9, 第78頁] 問題152。

9. 酉矩陣的遍歷定理

如果複數 u 滿足 $|u| \leq 1$, 則平均值

$$a_n = \frac{1}{n} \sum_{k=0}^{n-1} u^k$$

構成一個收斂的序列, 這是經典分析中一個簡單而有趣的結果。我們來看一下其證明。若 $u = 1$,

⁷參見 Gauss, *Gesammelta Werke*, Vol.10, 第24–25頁。英文引自 A. Borel, *Mathematics: art and science*, The Mathematical Intelligence, 5:4 (1983) 9–17. 中譯文《數學-藝術與科學》, 江嘉禾譯, 《數學譯林》第4卷 (1985年) 第3期, 243–253。

⁸矩陣 A 稱為正規矩陣, 如果 A 與 A^* 可交換。

則 $a_n = 1$, 所以 $\{a_n\}$ 收斂到 1。如果 $u \neq 1$, 則

$$a_n = \frac{1}{n} \sum_{k=0}^{n-1} u^k = \frac{1}{n} \frac{1-u^n}{1-u},$$

由於 $|u| \leq 1$, 所以 $|1-u^n| \leq 1+|u^n| \leq 1+1=2$, 從而

$$|a_n| = \frac{1}{n} \frac{|1-u^n|}{|1-u|} \leq \frac{1}{n} \frac{2}{|1-u|},$$

所以 $\{a_n\}$ 收斂到 0。

利用酉矩陣的譜定理 (每一個酉矩陣酉相似於一個對角矩陣), 可以將這一結果從複數 (一階矩陣) 推廣到高階矩陣上, 特別的, 我們有下述有名的 von Neumann 平均遍歷定理 (見 [8, 第 185 頁])。

定理 12: 設 U 是有限維內積空間 H 上的酉變換, 設 V 是 U 的不動點構成的子空間, 則由

$$A_n = \frac{1}{n}(I + U + \cdots + U^{n-1})$$

定義的序列當 $n \rightarrow \infty$ 時收斂到正交投影 P_V 。

這裏所謂“ A_n 收斂到 P_V ”是指, 對任意的 $x \in H$, $\|A_n x - P_V x\| \rightarrow 0$ 。

現在 von Neumann 本人的原始證明已經難得見到了, 因為文獻中通常收入的都是 F. Riesz 對這一定理給出的簡單而漂亮的證明, 這裏我們概述如下, 具體的細節讀者可以自行補充, 或者參見 [8], [9] 或 [22]。

證明: 設 R 為 $I-U$ 的值域, 則容易證明, 對 $x \in R$, $A_n x$ 收斂到 0。另一方面, 當 $x \in V$ 時, $A_n x = x$ 收斂到 x 。最後, 注意到 V 和 R 互為正交補。證畢。

事實上, F. Riesz 還進一步注意到, 上述證明可以推廣到所謂的壓縮算子, 即滿足 $\|T\| \leq 1$ 的算子, 其要點在於, 下面的引理成立。

引理 6: 設內積空間 H 上的算子 T 滿足對任意 $x \in H$, 有 $\|Tx\| \leq \|x\|$, 則 T 與 T^* 有相同的不動點。

對此有興趣的讀者, 我們推薦讀者去參看 [9] 或 [22]。

另一方面, 在矩陣的情形, 我們有下述最一般的結論 (見 [5, 第 560 頁]):

定理 13: 矩陣 C 使得

$$\frac{1}{n}(I + C + \cdots + C^{n-1})$$

收斂當且僅當 C 的所有特徵值的模長小於等於 1, 並且對於模長為 1 的特徵值, 其代數重數等於幾何重數。

10. 辛矩陣的行列式為 1 : Siegel 的證明

令域 F 上的 $2n$ 階矩陣

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix},$$

則 F 上滿足 $PJP' = J$ 的 $2n$ 階矩陣 P 稱為辛矩陣, F 上的所有 $2n$ 階辛矩陣構成一個群, 稱為辛群, 記為 $Sp(2n, F)$ 。辛群與正交群一起構成兩類最重要的典型群⁹。從某種程度上講, 辛矩陣要比正交矩陣簡單 (正如反對稱雙線性型要比對稱雙線性型簡單), 例如辛矩陣的行列式總是 1 (而正交矩陣的行列式則可以是 1 或 -1)。我們將對 $F = \mathbb{R}$ 的情形證明這一點, 方法來自 C. L. Siegel [25]。

定理 14: 實辛矩陣的行列式為 1。

證明: 設 $P \in Sp(2n, \mathbb{R})$, 寫成與 J 一致的分塊

$$P = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

於是 P 滿足 $AB' = BA'$, $AD' - BC' = I$, $CD' = DC'$ 。注意到, 當 $P \in Sp(2n, \mathbb{R})$ 時, $P' \in Sp(2n, \mathbb{R})$, 從而有 $A'C = C'A$, $A'D - C'B = I$, $B'D = D'B$ 。於是

$$\begin{pmatrix} iC' + D' & -iA' - B' \\ 0 & I \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & iI \\ 0 & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ C & iC + D \end{pmatrix},$$

兩邊取行列式有 $(\det P - 1) \det(iC + D) = 0$, 只要證明 $\det(iC + D) \neq 0$ 。注意到 $(iC + D)(-iC' + D') = CC' + DD'$, 只要證明 $CC' + DD'$ 可逆。用反證法, 不然, 假設存在非零向量 $x \in \mathbb{R}^n$ 使得 $(CC' + DD')x = 0$, 由此 (僅在此處用到基域 $F = \mathbb{R}$ 的假定) 推出 $C'x = D'x = 0$, 從而

$$\begin{pmatrix} A' & C' \\ B' & D' \end{pmatrix} \begin{pmatrix} 0 \\ x \end{pmatrix} = \begin{pmatrix} C'x \\ D'x \end{pmatrix} = 0,$$

此與 $P' \in Sp(2n, \mathbb{R})$ 可逆矛盾! 證畢。

⁹根據定理 4, 線性空間上最有意義的雙線性度量或者是對稱的或者是反對稱的, 所對應的保度量變換群分別是正交群與辛群。

或許有些奇怪，一個實數域的問題最終要放到複數域來解決，然而這並沒有什麼奇怪的，J. Hadamard 早就告誡過我們：“實數域中兩個真理之間的最短路程是通過複數域。”這在代數上就體現為域擴張的思想。這是一個常用的手段。正如 H. Weyl 所說的¹⁰：

如果在任意一個抽象的域而不是複數域中考慮問題，那麼此時代數基本定理不一定成立。該定理斷言，每個複變量的多項式可唯一分解為線性因子。因此，在代數研究中有一種習慣：看看一個證明是否用了代數基本定理。在每一種代數理論中，有一些屬於更基本的部分，它與代數基本定理無關，因此在所有的域中都成立；而對一些高深的部分，代數基本定理則是不可或缺的。對於這一部分，就需要有域的代數閉包。在大多數情況下，代數基本定理標誌著一種起決定作用的分界線；只要有可能就應該避免使用它。為建立在任意域中都成立的定理，將一個域嵌入到一個較大的域中的做法常常是有用的。特別地，有可能將任意一個域嵌入到一個代數閉域中。一個眾所周知的例子是一個實係數多項式在實數範圍內可分解為線性因子或二次因子。為了證明它，我們添加一個 i 到實數中，這樣便嵌入到複數域中了。這種方法在拓撲中有一個類比，用於對流形的研究與特性刻劃：在曲面的情形，這種類比就是應用覆蓋曲面。

11. 三元二次型的 Albert-Dickson 泛性定理

1929年，Dickson 證明了數論中三元二次型的泛性定理，這一結果曾經被 G. D. Birkhoff 列入美國數學會自 1878 年建立近 50 年以來美國數論研究最有代表性的成就。¹¹ 1933 年，Albert [1] 指出，Dickson 的數論定理事實上有一個代數的版本，也就是說，可以在任意的特徵不等於 2 的域 F (而不再是整數環 \mathbb{Z}) 上敘述這個結果，這就得到更一般的三元二次型的泛性定理。為了敘述這個定理，我們先引入幾個術語。我們稱 F 上向量空間 V 上的二次型 $Q(x)$ 能表示出 $a \in F$ ，如果存在 $x \in V$ 使得 $Q(x) = a$ ；稱 $Q(x)$ 是泛性的 (universal)，如果 $Q(x)$ 能表示出 F 的所有元素；稱向量 v 是 $Q(x)$ 的零化向量 (null vector)，如果 $Q(v) = 0$ 且 $v \neq 0$ 。並且，我們用 $[a_1, a_2, \dots, a_n]$ 表示二次型 $a_1x_1^2 + a_2x_2^2 + \dots + a_nx_n^2$ 。回憶起兩個二次型 $f(x)$ 與 $g(x)$ 稱為等價的，如果存在 V 上的可逆線性變換 T 使得 $g(x) = f(T(x))$ 。二次型 $Q(x) = x'Ax$ 的判別式 d 定義為對應的對稱矩陣 A 的行列式： $d = \det A$ 。於是，Albert 發現的泛性定理可以表述如下。

¹⁰H. Weyl, *Topology and abstract algebra as two Roads of mathematical comprehension* Part I. Amer. Math. Monthly, **102** (1995), 453-460, Part II. Amer. Math. Monthly, **102** (1995), 646-651. 中譯文《拓撲和抽象代數：理解數學的兩種途徑》，馮緒寧譯，數學譯林》第 5 卷 (1986 年) 第 1 期，74-87.

¹¹參見 G. D. Birkhoff 在 1938 年所作的演講 *Fifty years of American mathematics*，收入 *Semicentennial Addresses of the American Mathematical Society*, Vol.1, American Mathematical Society, 1980.

定理 15: 對於一個具有判別式 $d \neq 0$ 的三元二次型 $Q(x)$, 以下四個條件等價:

- (i) $Q(x)$ 有零化向量。
- (ii) $Q(x)$ 能表示出 $-d$ 。
- (iii) $Q(x)$ 是泛性的。
- (iv) $Q(x)$ 等價於 $[-d, 1, -1]$ 。

作為這個定理的一個推論, 我們考慮 $Q(x) = x_1^2 + x_2^2 + x_3^2$, 則我們得到以下有趣結果 (見 Kaplansky [13, 第 21 頁] 練習 10): 在一個特徵不等於 2 的域中, 如果 -1 可以寫成三個元素的平方和, 則 -1 也可以寫成兩個元素的平方和。(這就是從 (ii) \implies (i)。)

Albert 給出的原始證明是計算性質的, 有一點麻煩, 這裏我們給出一個更加簡單而清晰的幾何證明。這個證明是在 Kaplansky [13] 的啓發下完成的。我們將這個定理分解為以下三個引理。

引理 7: 設 $Q(x)$ 是一個具有判別式 $d \neq 0$ 的三元二次型, 則 $Q(x)$ 能表示出 $-d$ 當且僅當它等價於 $[-d, 1, -1]$ 。

引理 8: 設 $Q(x)$ 是一個具有判別式 $d \neq 0$ 的三元二次型, 則 $Q(x)$ 是泛性的當且僅當它等價於 $[-d, 1, -1]$ 。

引理 9: 設 $Q(x)$ 是一個具有判別式 $d \neq 0$ 的三元二次型, 則 $Q(x)$ 有零化向量當且僅當它等價於 $[-d, 1, -1]$ 。

這裏我們只給出引理 7 的證明, 引理 8 是它的簡單推論 (注意到二元二次型 $[1, -1]$ 在特徵不等於 2 的域 F 上是泛性的: 對任意的 $a \in F$ 有, $a = (\frac{a+1}{2})^2 - (\frac{a-1}{2})^2$), 而引理 9 是 Kaplansky [13, 第 20 頁] 的練習 3, 我們留給讀者。

引理 7 的證明:

證明: 根據熟知的對角化過程, 我們可以從 $-d$ 的表示向量 v 出發對角化 $Q(x)$, 假設我們最後得到的二次型是 $Q'(y) = [-d, a, b]$, 那麼根據相合關係下判別式相差一個非零平方因子的事實可知, $d = -dab \cdot c^2$, 即 $1 = (-1)abc^2$, 通過一個簡單的線性替換 $y_1 = z_1, y_2 = z_2, y_3 = cz_3$, 我們得到 $Q'(y) = [-d, a, b]$ 進一步等價於 $Q''(z) = [-d, a, -a^{-1}]$, 最後我們只需要對 $[a, -a^{-1}]$ 應用下述引理, 即可得到結論。證畢。

引理 10: 設 $Q(x)$ 是特徵不等於 2 的域上的一個非退化的二元二次型, 假定 $Q(x)$ 存在一個零化向量, 則 $Q(x)$ 等價於對角型 $[1, -1]$ 。

引理 10 就是 Kaplansky [13, 第 19 頁] 的定理 16, 這裏不再證明。爲了在引理 7 的證明中應用這一結果, 我們只需注意到 $a1^2 - a^{-1}a^2 = 0$, 即 $v = (1, a)'$ 是 $[a, -a^{-1}]$ 的零化向量。當然, 我們也完全可以不引用這個一般結果, 因爲事實上, $ax^2 - a^{-1}y^2$ 在線性替換 $x = z + a^{-1}w, y = az - w$ 下化爲 $4zw$, 它顯然等價於 $[1, -1]$ 。這裏之所以要給出引理 10, 不僅因爲它本身很有意思, 富有幾何上的趣味; 而且, 更重要的是, 它與下面將要提到的 Witt 消去定理一起構成了 Witt 分解定理 (見 Kaplansky [13, 第 20 頁] 的定理 18) 的基礎。

讀者如果拿上面的證明與 Albert 原來的證明比較, 很容易就會看出上述方法的優越性, 事實上, 這也正是幾何觀點優於代數觀點的體現。可惜的是, 在通常的教科書中, 對二次型的處理幾乎都是純代數的, 對幾何的觀點幾乎絕口不提, 只有 Kaplansky [13] 是一個例外。然而事實上, 二次型理論中很多結果的最好的表述與證明都需要採用幾何的語言, 下面我們再舉一個例子來說明這一點。

12. Witt 定理

這就是二次型的代數理論中具有基本重要性的 Witt 定理, 我們採用以下的表述:

定理 16: 在一個特徵不等於 2 的域上, 對一個非奇異的二次型 h , 如果 $h \oplus f$ 等價於 $h \oplus g$, 則 f 等價於 g 。

證明: 爲了證明 Witt 消去律, 只需要考慮 h 是一維型的情形; 換言之, 只需要證明, 如果 $[a, b, c, \dots]$ 等價於 $[a, b', c', \dots]$, 且 $a \neq 0$, 那麼 $[b, c, \dots]$ 等價於 $[b', c', \dots]$ 。

第一個等價條件可以表述爲: 空間存在一組正交基 e_1, e_2, e_3, \dots 滿足

$$(e_1, e_1) = a, \quad (e_2, e_2) = b, \quad (e_3, e_3) = c, \quad \dots,$$

同時存在另一組正交基 e'_1, e'_2, e'_3, \dots 滿足

$$(e'_1, e'_1) = a, \quad (e'_2, e'_2) = b', \quad (e'_3, e'_3) = c', \quad \dots$$

我們只需要找到這個空間的一個對稱將 e_1 變換到 e'_1 , 因爲它必定把正交於 e_1 的子空間 $\langle e_2, e_3, \dots \rangle$ 變換到正交於 e'_1 的子空間 $\langle e'_2, e'_3, \dots \rangle$ 。¹²現在關於垂直於向量 r 的超平面的反射由公式

$$x \mapsto x - \frac{2(x, r)}{(r, r)}r$$

給出。很容易驗證上述反射保持二次型 (x, x) 。如果取 $r = e_1 - e'_1$, 這個反射確實將 e_1 變換到 e'_1 。但是, 這裏存在一個問題: 對於 (r, r) 等於零的向量 r , 反射沒有定義。現在我們的

¹²這就是第二個等式的含義。

二次型 (x, x) 可能是一個不定的二次型，所以 $(e_1 - e'_1, e_1 - e'_1)$ 有可能等於零。但是，如果 $(e_1 - e'_1, e_1 - e'_1) = 0$ ，我們可以用向量 $s = e_1 + e'_1$ 作反射，它將 e_1 變換到 $-e'_1$ ，然後再對全空間取 -1 決定的相似。注意到， (r, r) 與 (s, s) 不可能都是為零，這是因為，根據 Apollonian 等式，我們有

$$(r, r) + (s, s) = (e_1 - e'_1, e_1 - e'_1) + (e_1 + e'_1, e_1 + e'_1) = 2(e_1, e_1) + 2(e'_1, e'_1) = 4a \neq 0.$$

證畢。

事實上，Witt 定理這一基本結果直到 1936 年才被 Witt 得到。對此，Artin 曾作出一個極有教育意義的評論 (引自 [3, 第二卷第 106 頁]):

Artin 把 Witt 定理稱為一個令人感到丟臉的定理，他的意思是說，一個無論從其陳述還是從其數學對象來看都是如此簡單的定理居然要等到 1936 年才被發現並證明，何況這個定理在各個領域都是極其有用的。

最近，Scharlau [24] 指出，事實上，Dickson 在 1907 年的文章裏就已經得到了 Witt 定理，不過他的這篇文章被完全忽略了。原因只在於 Dickson 的文章過分代數化了，讓人很難念，Scharlau 在 [24, 第 237–239 頁] 這樣寫到:

看來 Dickson 的文章好像被完全忽略了，我找不到一篇引用這個工作的文獻。然而，你必須承認，正如 Dickson 的許多其他文章一樣，這篇文章非常不好讀。它完全是代數的。對於 Witt (1937 年) 如此優美地應用的二次型的幾何解釋，他毫無所知。原本可以用一個簡單的幾何論證的地方，他卻用了冗長的代數計算。他沒有認識到反射的重要性。當他需要構造一個適當的等距時他利用 Cayley 變換，一個用起來並不方便的工具。然而，無論如何，我認為只要是考慮 Witt 定理或相關的問題，Dickson 的貢獻應該得到承認。

註: Witt 定理還有下面一個版本，常常稱之為 Witt 延拓定理 (見 Artin [2, 第 121 頁]):

定理 17: 設 Q 是有限維向量空間 V 上的非退化二次型， W_1, W_2 是 V 的子空間，且 $\sigma : W_1 \rightarrow W_2$ 是一個保持 Q 不變的線性滿射，則 σ 可以延拓成爲 V 上的等距變換。

13. 一個練習: Pauli 矩陣

筆者的論文 [16] 實際上是從一個習題開始的，或許這個小問題是值得了解的¹³

¹³事實上，筆者最初對這個問題感興趣，也是從一個習題引起的，見 F. W. Byron and R. W. Fuller, *Mathematics of Classic and Quantum Physics*, vol.1, Addison-Wesley Publishing Company (1968), p. 139, Ex.27. 中譯本《物理學中的數學方法》，熊家炯、曹小平譯，科學出版社，1984 年。

練習: 設 2 階複矩陣 A_1, \dots, A_k 滿足矩陣方程

$$\begin{cases} A_i^2 = I \\ A_i A_j = -A_j A_i \quad (i \neq j) \end{cases} \quad (i, j = 1, \dots, k),$$

證明

- (1) k 的最大值是 3, 並且這 3 個矩陣可以全部取為 Hermite 矩陣。
- (2) 若 A_1, \dots, A_k 都是對稱矩陣, 則 k 的最大值是 2。
- (3) 若 A_1, \dots, A_k 都是實矩陣, 則 k 的最大值是 2。

註: 在理論物理中, 著名的 Pauli 矩陣

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

滿足上述方程。Pauli 矩陣或許是理論物理中出現最為頻繁的矩陣, 原因正如費曼所一語道破的: “Pauli 的自旋矩陣和算子不是什麼新的東西, 而正是 Hamilton 的四元數”。¹⁴ (嚴格說來, Pauli 矩陣與 Hamilton 四元數的基矩陣 (見 [8, 第 69 頁] 習題 6) 之間相差一個因子 i 。) 複數的重要性在數學和物理學中已經得到了廣泛的認識, 而四元數的重要性則似乎有待進一步挖掘, 楊振寧先生在他的論文選中如是說¹⁵:

SU(2) 對稱的存在, 一定有一個理由, 因為在最根本的層次上, 造化的安排一定不會是無緣無故的, 這種說法已經不止一次得到應驗。除此之外, 我們期待著的解釋, 可能要用到四元數代數, 因為四元數的對稱就是 *SU(2)*。此外, 四元數代數是一種美麗的結構。雖然它不是交換的。但是我們知道, 造化選擇非交換的代數作為量子力學的語言, 她怎麼會拒絕使用這僅有的另一種可能的美妙代數作為她在宇宙萬物中建立起來的所有複雜對稱性的語言呢?

小結: 幾何與代數

總之, 為了使大部分學生能最有效地學習線性代數, 我們應該盡可能地強調幾何的語言和方法。幾何的語言, 自然是相對於代數的語言而說的。簡單地講, 就是用線性變換代替矩陣, 用抽象向量代替 n 維列向量。幾何語言的優點是簡潔明快, 例如“作用”這個詞給人的感覺就是如

¹⁴ 見 *The Pleasure of Finding Things Out: The Best Short Works of Richard P. Feynman*, edited by Jeffrey Robbins, Perseus Books, 1999. 中譯本《發現的樂趣》第 205 頁, 張鬱乎譯, 湖南科學技術出版社, 2007 年。

¹⁵ 見楊振寧, *Selected Papers 1945–1980 With Commentary*, San Francisco: Freeman & Co, 1983. 中譯文《優雅的四元數》, 收入《楊振寧文集》第 35–38 頁, 張奠宙主編, 華東師範大學出版社, 1998 年。

此。代數語言的好處是具體清晰，兩個矩陣“相乘”在我們頭腦中的圖像自然是一系列具體運算的運作。通常的教科書都過分強調了代數的語言，這同時也充分暴露了它的諸多缺點。最大的缺點在於坐標不具有內蘊的含義。或許可以提一句，正如陳省身先生所注意到的，愛因斯坦將狹義相對論推廣到廣義相對論花了7年之久就是因為他一直擺脫不了坐標的束縛¹⁶：

如果你覺得接受一般的坐標概念有困難，那麼你有一個好的夥伴。愛因斯坦花了7年時間才從狹義相對論過度到廣義相對論。他對之所以延遲了這麼久的解釋是：“為什麼建立廣義相對論又用了7年時間呢？主要原因是：要擺脫「坐標必須具有直接的度量意義」這個舊觀念是不容易的。”

在很多問題中坐標的選取並不重要，我們所需要的往往只是一些基本的運算規律，例如分對稱性、雙線性等。這時候抽象的幾何語言就十分適用了。¹⁷ 例如在內積空間的理論中，往往採用幾何的語言，而且事實上，線性代數的幾何直觀性在這裏得到了最大的體現。特別地，實對稱矩陣的譜定理可以敘述為：在歐氏空間中，對給定的對稱變換，存在一組由特徵向量構成的標準正交基。我們正是靠這種幾何觀點來指引具體的代數運算的，例如所謂 Schmidt 正交化，無非就是將第二個向量沿第一個向量作垂線，一旦指出這一點，Schmidt 正交化公式就很容易理解了。

本文主要強調了線性代數的幾何觀點，這是 von Neumann–Halmos 傳統的延續。用幾何的語言敘述的線性代數教科書，這裏向讀者推薦 Halmos [8]。但是應該指出，線性代數還有許多不同的側面。我們最後再介紹另外兩個不同的觀點作為補充，或許某些讀者是感興趣的。

Taussky 在 [27] 中敘述了她所了解的矩陣論，這也是一家之言，特別的，她在結語部分對讀者所提的幾點忠告很值得重提：

當你觀察到數的一個有趣性質時，也許你沒有把它當做是 $n \times n$ 矩陣的有趣性質在 $n = 1$ 的情形。請想一想， $GL(n, F)$ 或 $SL(n, F)$ ， $GL(n, \mathbb{Z})$ 或 $SL(n, \mathbb{Z})$ 。

當你有一對有趣性質的矩陣時，請研究它們生成的束或代數。

當某一個矩陣的行列式被證明是重要的時候，試從整體上探究一下這個矩陣，例如，是否可以作為代數數域的判別式矩陣。

當某一個一元多項式使你感興趣時，試考慮一下以之為特徵多項式的矩陣。

當有人瞧不起矩陣時，請回想一下對矩陣論作出了重要貢獻的大數學家，比如，Frobenius, Schur, Siegel, Ostrowski, Motzkin, Kac, I. M. Gelfand 等。

此外，筆者在2010年參加了在北京清華大學召開的第五屆世界華人數學家大會，聆聽了 Gilbert Strang 所做的報告，按照他的說法，Israel Gohberg(1928–2009) 是近一百年裏最偉

¹⁶見陳省身《從三角形到流形》，收入《陳省身文集》pp. 233–243，張奠宙主編，華東師範大學出版社，2002年。

¹⁷所以，H. Weyl 說，“以坐標的形式把數引進幾何學，是一種暴力行為”。見上一註腳中提到的陳省身的文章第 237 頁。

大的線性代數專家，對此筆者也不甚理解，或許是因為 Gohberg 的工作更偏向於應用吧¹⁸。

致謝

我總結這份材料是響應 2006 年徐澤同學的一個建議，她認為在畢業時可以給下一屆的學弟學妹們留一點什麼，或許對他們有幫助。這就是本文的初衷。

其次，我要感謝劉雲朋同學，這裏的許多數學點子都蒙他賜教。感謝首都師範大學趙潔、段紅偉同學，作者從與她們的討論中獲益良多。

最後，我要感謝我的線性代數老師田代軍，他向我推薦了 Jacobson 的經典教科書[12]以及華羅庚與萬哲先合著的《典型群》，讓我體會到學習線性代數的快樂。

特別要感謝審稿人，對作者的初稿提出了許多有益的建議。

在本文的準備過程中，作者還得到了北京清華大學數學系周堅教授、北京大學數學系丘維聲教授、上海復旦大學數學系姚慕生教授、臺灣國立交通大學應用數學系吳培元教授以及《數學傳播》編輯部陳麗伍助理的慷慨幫助，在此一併表示感謝。

參考資料

1. A. A. Albert, *A note on the Dickson theorem on universal ternaries*, Bull. Amer. Math. Soc., **39**(1933), 585–588.
2. E. Artin, *Geometric Algebra*, Interscience, New York, 1957.
3. M. Berger, *Geometry*, (Corrected fourth printing), Springer, 2009. 中譯本《幾何》，陳志杰、周克希譯，科學出版社，1989年。
4. G. W. Dunnington, *Gauss: Titan of Science*, Math. Asso. Amer., 2004.
5. N. Dunford and J.T. Schwartz, *Linear Operators, Part I, General Theory*, John Wiley & Sons, 1988.
6. H. Flanders, *Elementary divisors of AB and BA* , Proc. Amer. Math. Soc., **2**(1951), 871–874.
7. P. R. Halmos, *Hilbert Space and the Theory of Spectral Multiplicity*, Chelsea, 1951.
8. P. R. Halmos, *Finite Dimensional Vector Spaces*, Springer, 1974.
9. P. R. Halmos, *A Hilbert Space Problem Book*, Springer-Verlag, 1967. 中譯本《希爾伯特空間問題集》，林辰譯，上海科學技術出版社，1984年。
10. P. R. Halmos, *I Want to Be a Mathematician*, Springer-Verlag, 1985. 中譯本《我要做數學家》，馬元德、沈永歡、胡作玄、趙慧琪譯，江西教育出版社，1999年。
11. P. R. Halmos, *Linear Algebra Problem Book*, Math. Asso. Amer., 1995.
12. N. Jacobson, *Lectures in Abstract Algebra II. Linear Algebra*, GTM31, 1953. 中譯本

¹⁸最近筆者就這個問題請教了 MIT 數學系的 G. Strang 教授 (因為授課廣受學生好評，他給人的印象大概就是 MIT 的線性代數“教父”吧)，他答覆如下：一個方法是從亞馬遜網頁上流覽一下他的著作標題。而且現在有許多奉獻給他的書，還有許多線性代數方面的論文。他與 Harry Dym 解決的一個小問題如下：如果你有一個三對角的實對陣矩陣，主對角元占優，如何將它補充成一個實對稱矩陣，使它的行列式取得最大值？補充之後的實對稱矩陣的逆又是什麼？它為什麼是三對角的？也許你可以讓學生求解一個 3×3 的例子。(大概是我在通信中沒有說清楚自己的學生身份，以至於讓 Strang 教授誤以為我是一名教師了！)

- 《抽象代數學 卷2 線性代數》，黃緣芳譯，科學出版社，1960年。
13. I. Kaplansky, *Linear Algebra and Geometry-A Second Course*, Allyn and Bacon, 1974.
 14. E. Landau, *Elementary Number Theory*, AMS Chelsea Publishing, 1999.
 15. 林開亮, Hua 引理及其應用, 《數學傳播》第34卷 (2010年) 第135期, 39–48。
 16. 林開亮, Hurwitz-Radon 矩陣方程, 《數學傳播》, 第38卷 (2012年) 第141期, 48–63。
 17. L. H. Loomis and S. Sternberg, *Advanced Calculus*, Addison-Wesley Publishing Company, 1968. 中譯本《高等微積分》, 王元、胥鳴偉譯, 高等教育出版社, 2006年。
 18. L. Lovász, *Combinatorial Problems and Exercises*, Second Edition, AMS Chelsea Publishing, 2007.
 19. J. von Neumann, *Distribution of the ratio of the mean square successive difference to the variance*, Ann. Math. Stat., **12**(1941), 367–395. (收入 *Collected Works*, Vol.4, 452–480, 特別的, 見第456–457頁。)
 20. J. von Neumann, *Approximative properties of matrices of higher finite order*, Portugaliae Math.**3**(1942), 1–62. (收入 *Collected Works*, Vol.4, 270–331, 特別的, 見第330頁。)
 21. R. Remmert, *Theory of Complex Analysis*, Graduate Texts in Mathematics, Reading in Mathematics, V.122, Springer, 1991.
 22. F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Dover Publications Inc, New York, 1990.
 23. 丘維聲, 《高等代數》(下冊), 清華大學出版社, 2010年。
 24. W. Scharlau, *On the history of the algebraic theory of quadratic forms*, 收入 *Quadratic Forms and Their Applications*, Contemporary Mathematics **272** A.M.S., 2006.
 25. C. L. Siegel, *Symplectic geometry*, Amer. J. Math., **65**(1943), 1–86.
 26. M. Spivak, *Calculus on Manifold*, Benjamin, New York, 1965. 中譯本《流形上的微積分》, 齊民友、路見可譯, 人民郵電出版社, 2005年。
 27. O. Taussky, *How I became a torchbearer for matrix theory*, Amer. Math. Monthly, **95** (1988), 801–812. 我是怎樣成爲矩陣論傳人的, 馮慈黃譯, 《數學譯林》第9卷 (1990年) 第2期, 119–129.
 28. J.-H. Wang and P. Y. Wu, *Difference and similarity models of two idempotent operators*, Linear Algebra Appl., 208/209 (1994), 257–282.
 29. W. C. Waterhouse, *Cochran's theorem for rings*, Amer. Math. Monthly, **108** (2001), 58–59.
 30. 姚慕生、吳泉水, 《高等代數學》, 復旦大學出版社, 2008年。
 31. 張堯庭, 深深的懷念—我所知道的許寶騷先生, 收入《道德文章垂範人間: 紀念許寶騷先生百年誕辰》, 344–356, 北京大學出版社, 2010年。

—本文作者爲中國首都師範大學數學科學院研究生—