線代啟示錄

I seek not to know the answers, but to understand the questions.

矩陣函數 (上)

Posted on 12/28/2010 by ccjou

本文的閱讀等級：中級

給定二階方陣 $A=\begin{bmatrix}-2&-6\\1&3\end{bmatrix}$，如何計算 $\cos A$？(取自交大資訊所2007年入學試題) 我們或許直覺認為 $\cos A$ 各元不過就是 $A=[a_{ij}]$ 對應元的餘弦函數，$(\cos A)_{ij}=\cos a_{ij}$，上例為

$$\cos\begin{bmatrix}-2&-6\\1&3\end{bmatrix}=\begin{bmatrix}\cos(-2)&\cos(-6)\\\cos 1&\cos 3\end{bmatrix}。$$

此定義的缺點在於 $\cos A$ 未能保留餘弦函數的一些美好性質。舉例而言，既然有 $\cos^2\theta+\sin^2\theta=1$ 和倍角公式 $\cos(2\theta)=2\cos^2\theta-1$，我們自然希望任意方陣 A 的矩陣函數 $\cos A$ 和 $\sin A$ 同樣滿足 $\cos^2 A+\sin^2 A=I$ 和 $\cos(2A)=2\cos^2 A-I$。但這要如何辦到呢？本文僅解說可對角化矩陣函數，不可對角化矩陣函數涉及 Jordan 形式，將留待下文詳細討論。

給定一 $n\times n$ 階矩陣 A，目前存在許多種矩陣函數定義方法，而基於線積分 (line integral) 的定義最被多數人採用[1]，但其中涉及了複變函數，在此我們採用另一種便捷的導入方式。我們使用讀者熟悉的矩陣指數 (見“矩陣指數”) 來定義一般矩陣函數 $f(A)$。為了讓矩陣函數與同型純量函數擁有相同的性質，我們從泰勒展開式出發，以指數函數為例，

$$e^x=1+x+\frac{x^2}{2!}+\cdots。$$

將 x 以 A 取代，常數 1 以單位矩陣 I 取代 (因為 $x^0=1$，也就有 $A^0=I$)，便得到矩陣指數

$$e^A=I+A+\frac{A^2}{2!}+\cdots。$$

無窮級數衍生出收斂性問題，最直接的解決方式是從 e^A 的表達式切入。若 A 可對角化為

$$A=S\Lambda S^{-1}=S\begin{bmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{bmatrix}S^{-1}，$$

其中 $\Lambda=\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$ 的主對角元 λ_j 即為 A 的特徵值，則冪矩陣 A^k 也可對角化為 $A^k=S\Lambda^kS^{-1}$，而 $\Lambda^k=\mathrm{diag}(\lambda_1^k,\ldots,\lambda_n^k)$。利用 A^k 的對角化式子化簡 e^A，可得

$$\begin{aligned}e^A&=SIS^{-1}+S\Lambda S^{-1}+S\frac{\Lambda^k}{2!}S^{-1}+\cdots\\&=S\left(I+\Lambda+\frac{\Lambda^2}{2!}+\cdots\right)S^{-1}\\&=S\begin{bmatrix}e^{\lambda_1}&&\\&\ddots&\\&&e^{\lambda_n}\end{bmatrix}S^{-1}.\end{aligned}$$

若將 A 替換為 Λ，則 $S=I$，也就有 $e^\Lambda=\mathrm{diag}(e^{\lambda_1},\ldots,e^{\lambda_n})$，故 e^A 也可以寫為 $e^A=Se^\Lambda S^{-1}$。將上述結果推廣至任意函數 $f(x)$，同樣可定義對角矩陣函數

$$f(\Lambda)\overset{\text{def}}{=}\mathrm{diag}(f(\lambda_1),\ldots,f(\lambda_n))，$$

並且令

$$f(A)=Sf(\Lambda)S^{-1}=S\begin{bmatrix}f(\lambda_1)&&\\&\ddots&\\&&f(\lambda_n)\end{bmatrix}S^{-1}。$$

表面上，從矩陣指數 e^A 推廣至一般矩陣函數 $f(A)$ 的過程似乎顯得粗略草率，我們如何確知由無窮級數定義出的 $f(A)$ 具備收斂性並且唯一存在？ 先說明收斂性。設 $f(z)$ 為一解析函數 (analytic function)，亦即對於 $|z-z_0|<\rho$，下

列冪級數收斂：

$$f(z)=\sum_{k=0}^{\infty}c_k(z-z_0)^k。$$

若可對角化矩陣 A 的每一特徵值 λ_i 都滿足 $|\lambda_i-z_0|<\rho$，可知 $f(\lambda_i)$ 收斂，推論 $f(\Lambda)$ 也收斂，因此確認 $f(A)=Sf(\Lambda)S^{-1}$ 具收斂性。

方陣 A 並沒有唯一的對角化形式 $S\Lambda S^{-1}$，特徵值矩陣 Λ 的主對角元可調換位置，此外，對應重複特徵值，我們也可以選擇不同的特徵向量組成 S。欲證明 $f(A)$ 唯一存在，我們藉助可對角化矩陣的譜分解表達式。設 $\{\lambda_1,\ldots,\lambda_m\}$ 代表 A 的相異特徵值集合，也稱為矩陣譜，並令對應 λ_i 的相重數為 β_i。令 $n\times\beta_i$ 階矩陣 X_i 的行向量由對應 λ_i 的特徵向量組成，亦即 $C(X_i)=N(A-\lambda_iI)$ 且 $\mathrm{dim}C(X_i)=\beta_i$，因此特徵向量矩陣 S 可以表示為 $S=\begin{bmatrix}X_1&\cdots&X_m\end{bmatrix}$，$S^{-1}$ 也可以寫為 $S^{-1}=\begin{bmatrix}Y_1^T\\ \vdots\\ Y_m^T\end{bmatrix}$，$Y_i^T$ 為 $\beta_i\times n$ 階分塊，則 $f(A)$ 的矩陣譜分解式如下：

$$\begin{aligned}f(A)&=S\Lambda S^{-1}\\&=\begin{bmatrix}X_1&\cdots&X_m\end{bmatrix}\begin{bmatrix}f(\lambda_1)I_{\beta_1}&&\\&\ddots&\\&&f(\lambda_m)I_{\beta_m}\end{bmatrix}\begin{bmatrix}Y_1^T\\ \vdots\\ Y_m^T\end{bmatrix}\\&=f(\lambda_1)X_1Y_1^T+\cdots+f(\lambda_m)X_mY_m^T\\&=f(\lambda_1)P_1+\cdots+f(\lambda_m)P_m,\end{aligned}$$

其中 $P_i=X_iY_i^T$ 即為沿著子空間 $C(A-\lambda_iI)$ 至特徵空間 $N(A-\lambda_iI)$ 的投影矩陣 (證明見“可對角化矩陣的譜分解”)，P_i 由 A 唯一決定。所以，不管 Λ 如何排列特徵值，以及 S 矩陣如何選擇特徵向量，我們可以推斷 $f(A)$ 唯一存在。

回到本文初給的二階方陣例子，$A=\begin{bmatrix}-2&-6\\1&3\end{bmatrix}$是不可逆矩陣，必定有特徵值 $\lambda_1=0$，利用跡數性質 $\mathrm{trace}A=\lambda_1+\lambda_2=(-2)+3=1$，可知另一特徵值為 $\lambda_2=1$。再計算對應的特徵向量，得到 $\mathbf{x}_1=\begin{bmatrix}3\\-1\end{bmatrix}$，$\mathbf{x}_2=\begin{bmatrix}-2\\1\end{bmatrix}$，$A$ 可對角化為

$$A=\begin{bmatrix}3&-2\\-1&1\end{bmatrix}\begin{bmatrix}0&0\\0&1\end{bmatrix}\begin{bmatrix}3&-2\\-1&1\end{bmatrix}^{-1},$$

故餘弦矩陣函數為

$$\begin{aligned}\cos A&=\begin{bmatrix}3&-2\\-1&1\end{bmatrix}\begin{bmatrix}\cos 0&0\\0&\cos 1\end{bmatrix}\begin{bmatrix}3&-2\\-1&1\end{bmatrix}^{-1}\\&=\begin{bmatrix}3-2\cos 1&6-6\cos 1\\-1+\cos 1&-2+3\cos 1\end{bmatrix}.\end{aligned}$$

最後補充一個 $n\times n$ 階可對角化矩陣 A 的投影矩陣 P_i 的算法，此法無須解出特徵向量。根據 Cayley-Hamilton 定理，冪矩陣 A^k 可表示為 $I,A,\ldots,A^{n-1}$ 的線性組合 (見“利用 Cayley-Hamilton 定理計算矩陣函數”)，因此前述矩陣函數 $f(A)$ 的無窮冪級數表達式可簡化為 A 的多項式。換句話說，若 $f(A)$ 存在，我們可以得到一多項式 $p(z)$ 使得 $p(A)=f(A)$。作法如下：如果 $p(\lambda_i)=f(\lambda_i)$，$1\le i\le m$，則譜分解式給出

$$p(A)=\sum_{i=1}^m p(\lambda_i)P_i=\sum_{i=1}^m f(\lambda_i)P_i=f(A)。$$

然而，是否總存在一多項式使得 $p(\lambda_i)=f(\lambda_i)$，$1\le i\le m$？是的。答案是 Lagrange 內插多項式 (見“線性世界的根基——疊加原理”)

$$p(z)=f(\lambda_1)\prod_{j\neq 1}\left(\frac{z-\lambda_j}{\lambda_1-\lambda_j}\right)+\cdots+f(\lambda_m)\prod_{j\neq m}\left(\frac{z-\lambda_j}{\lambda_m-\lambda_j}\right),$$

所以矩陣函數 $f(A)$ 即為

$$f(A)=p(A)=f(\lambda_1)\prod_{j\neq 1}\left(\frac{A-\lambda_jI}{\lambda_1-\lambda_j}\right)+\cdots+f(\lambda_m)\prod_{j\neq m}\left(\frac{A-\lambda_jI}{\lambda_m-\lambda_j}\right)。$$

以 $g_i(z)=\begin{cases}1&\text{if } z=\lambda_i\\0&\text{if } z\neq\lambda_i\end{cases}$ 取代 $f(z)$，可得

$$g_i(A)=P_i=\prod_{j\neq i}\left(\frac{A-\lambda_jI}{\lambda_i-\lambda_j}\right),\quad 1\le i\le m\text{。}$$

上例中，對於 $\lambda_1=0$ 和 $\lambda_2=1$，

$$P_1=\frac{A-1I}{0-1}=\begin{bmatrix}3&6\\-1&-2\end{bmatrix},\quad P_2=\frac{A-0I}{1-0}=\begin{bmatrix}-2&-6\\1&3\end{bmatrix}\text{。}$$

所以，

$$\begin{aligned}\cos A&=(\cos 0)\begin{bmatrix}3&6\\-1&-2\end{bmatrix}+(\cos 1)\begin{bmatrix}-2&-6\\1&3\end{bmatrix}\\&=\begin{bmatrix}3-2\cos 1&6-6\cos 1\\-1+\cos 1&-2+3\cos 1\end{bmatrix}.\end{aligned}$$

引用文獻：

[1] Gene H. Golub and Charles F. Van Loan, *Matrix Computations*, 2nd ed., 1989, pp. 540.

繼續閱讀：

- 矩陣函數 (下)

This entry was posted in 特徵分析, 主題專欄 and tagged 矩陣譜, 矩陣函數, 矩陣指數, Lagrange 內插多項式, 分塊矩陣, 對角化. Bookmark the permalink.

13 則回應給 矩陣函數 *(*上*)*

Watt Lin 說:

12/28/2010 at 8:07 下午

請問有沒有一種性質存在：
det(exp(A)) = exp(det(A))
det(cos(A)) = cos(det(A))
det(sin(A)) = sin(det(A))

抱歉，我沒正式修線性代數，只是看過一遍老師的光碟，當作業餘進修。
問這個題目，可能很幼稚。
如果 exp, cos, sin 不具有這種性質，
可否請老師舉些例子，說明某些矩陣可以先計算行列式值再代入函數，恰好與先求函數再算行列式值，其結果相等。

對角矩陣用在exp函數，這項性質應該成立吧！其他的我不清楚。

我好像曾經看過其他類似的東西，但記憶模糊，請老師指導。

回覆

Watt Lin 說:

12/28/2010 at 8:24 下午

自己簡單驗證：
對角矩陣先求exp再作det，與先求det再作exp，是不同的。

回覆

---

ccjou 說:

12/28/2010 at 9:22 下午

你寫的性質都不存在，不過卻有
$\det(e^A)=e^{\lambda_1}\cdots e^{\lambda_n}=e^{\mathrm{trace}A}$

又因為 $\cos A=S(\cos\Lambda)S^{-1}$，所以
$\det(\cos A)=(\det S)(\det(\cos\Lambda))(\det S^{-1})$

$=(\cos\lambda_1)\cdots(\cos\lambda_n)$
但上式並不等於 $\cos(\det A)=\cos(\lambda_1\cdots\lambda_n)$

回覆

---

ccjou 說:
12/28/2010 at 9:27 下午

oops, 不小心按到submit
若 A 可對角化，一般情況即為
$\det(f(A))=\det(Sf(\Lambda)S^{-1})$
$=\det(f(\Lambda))=f(\lambda_1)\cdots f(\lambda_n)$

回覆

Watt Lin 說:
12/29/2010 at 1:33 下午

感謝老師說明，我的觀念清楚了！

回覆

阿蟲 說:
03/16/2011 at 7:51 下午

√A這類的呢,還有A^(1/3)
沒有馬克勞林級數可以展(還是有我不知道,我也搞不懂)
那它要怎麼解?

回覆

---

ccjou 說:
03/17/2011 at 8:00 上午

按上文所述方式同樣可以定義 $\sqrt{A}$ 和 $(A)^{1/3}$，但通常我們僅針對半正定矩陣定義其平方根（因為它具有實際用途）。對於所有

實數 $x\ge 0$，我們定義 $f(x)=\sqrt{x}$ 為 x 的正平方根。若 A 是半正定矩陣，則 A 的所有特徵值都是非負數，故 $\sqrt{A}$ 也是半正定，而且可以證明它唯一存在，因此也為良好定義。

下面是維基百科的介紹：
http://en.wikipedia.org/wiki/Square_root_of_a_matrix

另外你當然也可以利用 Taylor 展開計算 $\sqrt{A}$，在 $x=1$ 展開 $f(x)=\sqrt{x}$ 可得
$\sqrt{x}=1+\frac{1}{2}(x-1)-\frac{1}{4(2!)}(x-1)^2+\frac{1\cdot 3}{8(3!)}(x-1)^3+\cdots$
$=1+\frac{1}{2}(x-1)+\sum_{n=2}^{\infty}(-1)^{n+1}\frac{(2n-3)!}{2^{2n-2}(n!)(n-2)!}(x-1)^n$
將 x 替換為 A，1 替換為 I (單位矩陣) 即可。

回覆

ccjou 說:
03/17/2011 at 1:15 下午

補充說明：矩陣函數 $\sqrt{A}$ 是一個很複雜的問題，由以下數個事實可略窺一二。
1) 若 A 可對角化，則 $\sqrt{A}$ 可如上文所述方式計算。
2) 某些 A 可能存在無窮多種可能的 $\sqrt{A}$，例如，$X^2=I$，$X=S\begin{bmatrix}1&0\\0&-1\end{bmatrix}S^{-1}$，$S$ 為任意可逆矩陣。
3) 某些 A 沒有 $\sqrt{A}$，例如，
$A=\begin{bmatrix}0&1\\0&0\end{bmatrix}$
4) $A=\begin{bmatrix}0&1&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{bmatrix}$不可對角化，但 $\sqrt{A}$ 可以為
$\sqrt{A}=\begin{bmatrix}0&0&1&0\\0&0&0&1\\0&1&0&0\\0&0&0&0\end{bmatrix}$

回覆

阿蟲 說:
03/19/2011 at 12:20 上午

非常謝謝老師的解說

回覆

張盛東 說:

04/20/2013 at 4:35 上午

老師，請教一條關於矩陣函數收斂的問題。

For an arbitrary $A\in C^{m\times m}$ and norm $\latex ||\cdot||$, prove using Theorem "Every square matrix A has a Schur factorization".

(a) $\latex lim_{n\rightarrow \infty}||A^n||=1 \Leftrightarrow \rho(A) < 1$, where $\latex \rho$ is the spectral radius.

(b) $\latex lim_{t\rightarrow \infty}||e^{tA}||=0 \Leftrightarrow \alpha(A) < 0$, where $\latex \alpha$ is the spectral abscissa.

回覆

張盛東 說:

04/20/2013 at 4:37 上午

不好意思，多加了斜杠。

For an arbitrary and norm $||\cdot||$, prove using Theorem "Every square matrix A has a Schur factorization".

(a) $lim_{n\rightarrow\infty}||A^n||=1\Leftrightarrow\rho(A)<1$, where ρ is the spectral radius.

(b) $lim_{t\rightarrow\infty}||e^{tA}||=0\Leftrightarrow\alpha(A)<0$, where α is the spectral abscissa.

回覆

ccjou 說:

04/20/2013 at 8:22 上午

(1) Schur factorization: 設 A 是 $m\times m$。$A=UTU^*=U(D+N)U^*$，其中 $UU^*=I$，$D=\text{diag}(\lambda_1,\ldots,\lambda_m)$，上三角形矩陣$N$ 的主對角元為零，N 是nilpotent matrix，即 $N^k=0$，$k\ge m$。

(2) $A^n=U(D+N)^nU^*$，$\|A^n\|\le\|U\|\cdot\|(D+N)^n\|\cdot\|U^*\|=\|(D+N)^n\|$。見 https://ccjou.wordpress.com/2012/07/03/%E8%AD%9C%E5%8D%8A%E5%BE%91%E8%88%87%E7%9F%A9%E9%99%A3%E7%AF

(3) 使用二項式定理，見
https://ccjou.wordpress.com/2012/07/04/%E5%BB%A3%E7%BE%A9%E6%94%B6%E6%96%82%E7%9F%A9%E9%99%A3/
或
https://ccjou.wordpress.com/2010/12/08/%E6%94%B6%E6%96%82%E7%9F%A9%E9%99%A3/
可證明 $\lim_{n\to\infty}(D+N)^n=0$ 即 $\lim_{n\to\infty}\Vert A^n\Vert=0$ iff $\rho(A)<1$。你寫出的命題有誤。第二個問題用類似方法亦可得證。

回覆

張盛東 說:
04/20/2013 at 1:16 下午

是的。第一問應該是$lim_{n\to\infty}||A^n||=0$。謝謝老師解答。

回覆