矩陣指數

維基百科，自由的百科全書

矩陣指數是方塊矩陣的一種矩陣函數，與指數函數類似。矩陣指數給出了矩陣李代數與對應的李群之間的關係。

設X為$n\times n$的實數或複數矩陣。X的指數，用e^X或exp(X)來表示，是由以下冪級數所給出的$n\times n$矩陣：

$$e^X = \sum_{k=0}^{\infty} \frac{1}{k!} X^k.$$

以上的級數總是收斂的，因此X的指數是定義良好的。注意，如果X是1×1的矩陣，則X的矩陣指數就是由X的元素的指數所組成的1×1矩陣。

目錄

性質

設X和Y為$n\times n$的複數矩陣，並設a和b為任意的複數。我們把$n\times n$的單位矩陣記為I，把零矩陣記為0。矩陣指數滿足以下性質：

- $e^0 = I$。
- $e^{aX}e^{bX} = e^{(a+b)X}$。
- $e^X e^{-X} = I$。
- 如果$XY = YX$，則$e^X e^Y = e^Y e^X = e^{(X+Y)}$。
- 如果Y是可逆的，則$e^{YXY^{-1}} = Ye^XY^{-1}$。
- $\det(e^X) = e^{\mathrm{tr}(X)}$，其中$\mathrm{tr}(X)$是$X$的跡數。
- $\exp(X^T) = (\exp X)^T$，其中X^T表示X的轉置。從中可以推出，如果X是對稱矩陣，則e^X也是對稱矩陣；如果X是斜對稱矩陣，則e^X是正交矩陣。
- $\exp(X^*) = (\exp X)^*$，其中X^*表示X的共軛轉置。可以推出，如果X是埃爾米特矩陣，則e^X也是埃爾米特矩陣；如果X是斜埃爾米特矩陣，則e^X是酉矩陣。

線性微分方程

矩陣指數的一個重要性，是它可以用來解微分方程。從(1)可知，以下微分方程

$$\frac{d}{dt}y(t) = Ay(t), \quad y(0) = y_0,$$

其中A是矩陣，具有解

$$y(t) = e^{At}y_0.$$

矩陣指數也可以用來解非齊次方程：

$$\frac{d}{dt}y(t) = Ay(t) + z(t), \quad y(0) = y_0.$$

參見以下的例子。

當A不是常數時，以下形式的微分方程沒有閉式解：

$$\frac{d}{dt}y(t) = A(t)\,y(t), \quad y(0) = y_0,$$

但馬格努斯級數可以給出無窮級數形式的解。

指數相加

我們知道，對於任何實數（純量）x和y，指數函數都滿足公式$e^{x+y} = e^x e^y$。類似的等式對於可交換矩陣也成立：如果矩陣X和Y是可交換的（意思是說$XY = YX$），則：

$$e^{X+Y} = e^X e^Y.$$

但是，如果它們不是可交換的，則以上的等式不一定成立。在這種情況下，我們可以用Baker-Campbell-Hausdorff公式來計算e^{X+Y}。

這個命題反過來不成立：$e^{X+Y} = e^X e^Y$並不一定就意味著X和Y是可交換的。但是，如果X和Y只含有代數數，而且它們的大小至少為2×2，則反過來也成立。（Horn & Johnson 1991，pp.435–437）.

指數映射

注意矩陣的指數總是非奇異方陣。e^X的逆矩陣由e^{-X}給出。這與複數的指數總是非零的事實類似。這樣，矩陣指數就給出了一個映射：

$$\exp\colon M_n(\mathbb{C}) \to \mathrm{GL}(n,\mathbb{C})$$

這是從所有$n\times n$矩陣的空間到一般線性群（所有非奇異方陣所組成的群）的映射。實際上，這個映射是滿射，就是說每一個非奇異方陣都可以寫成某個矩陣的指數。矩陣對數就是這個映射的逆映射。

對於任何兩個矩陣X和Y，我們有：

$$\|e^{X+Y} - e^X\| \le \|Y\| e^{\|X\|} e^{\|Y\|},$$

其中$\|\cdot\|$表示任意的矩陣範數。從中可以推出，指數映射在$M_n(\mathbf{C})$的緊子集內是連續和利普希茨連續的。

以下的映射

$$t \mapsto e^{tX}, \qquad t \in \mathbb{R}$$

定義了一般線性群中的一條光滑曲線，當$t = 0$時穿過單位元。實際上，這給出了一般線性群的一個單參數子群，這是由於：

$$e^{tX}e^{sX} = e^{(t+s)X}.$$

這條曲線在點t的導數（或切向量）由以下等式給出：

$$\frac{d}{dt}e^{tX} = Xe^{tX}. \qquad (1)$$

$t = 0$時的導數就是矩陣X，所以我們可以說，X是這個單參數子群的推廣。

更加一般地：

$$\frac{d}{dt}e^{X(t)} = \int_0^1 e^{(1-\alpha)X(t)}\,\frac{dX(t)}{dt}e^{\alpha X(t)}\,d\alpha.$$

矩陣指數的計算

尋找計算矩陣指數的可靠和準確的方法是困難的，目前在數學和數值分析領域中仍然是一個正在研究的話題。有些方法列舉如下。

可對角化矩陣

如果矩陣是對角的：

$$A = \begin{bmatrix} a_1 & 0 & \ldots & 0 \\ 0 & a_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_n \end{bmatrix},$$

則把主對角線上的所有元素取指數，就是原矩陣的指數：

$$e^A = \begin{bmatrix} e^{a_1} & 0 & \ldots & 0 \\ 0 & e^{a_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & e^{a_n} \end{bmatrix}.$$

這也允許了我們計算可對角化矩陣的指數。如果$A = UDU^{-1}$，且D是對角矩陣，則$e^A = Ue^DU^{-1}$。用西爾維斯特公式，也可以得到相同的結果。

冪零矩陣

如果對於某個整數q，有$N^q = 0$，則矩陣N稱為冪零矩陣。在這種情況下，矩陣指數e^N可以直接從級數展開式來計算，這是因為級數在有限個項後就終止了：

$$e^N = I + N + \frac{1}{2}N^2 + \frac{1}{6}N^3 + \cdots + \frac{1}{(q-1)!}N^{q-1}.$$

推廣

當矩陣X的最小多項式可以分解為一次多項式的積時，它就可以表示為以下的和：

$$X = A + N$$

其中：

- A是可對角化矩陣；
- N是冪零矩陣；
- A與N是可交換的（也就是說，$AN = NA$）。

這稱為Dunford分解。

這就是說，我們可以通過化為前兩種情況，來計算X的指數：

$$e^X = e^{A+N} = e^A e^N.$$

注意為了讓最後一步成立，A和N必須是可交換的。

另外一個密切相關的方法，是利用X的若爾當標準型。假設$X = PJP^{-1}$，其中J是X的若爾當標準型。那麼：

$$e^X = Pe^J P^{-1}.$$

另外，由於

$$J = J_{a_1}(\lambda_1) \oplus J_{a_2}(\lambda_2) \oplus \cdots \oplus J_{a_n}(\lambda_n),$$

$$\begin{aligned} e^J &= \exp\left(J_{a_1}(\lambda_1) \oplus J_{a_2}(\lambda_2) \oplus \cdots \oplus J_{a_n}(\lambda_n)\right) \\ &= \exp\left(J_{a_1}(\lambda_1)\right) \oplus \exp\left(J_{a_2}(\lambda_2)\right) \oplus \cdots \oplus \exp\left(J_{a_k}(\lambda_k)\right). \end{aligned}$$

因此，我們只需要知道怎樣計算若爾當塊的矩陣指數。但是，每一個若爾當塊都具有形式

$$J_a(\lambda) = \lambda I + N$$

其中N是冪零矩陣。則這個區塊的矩陣指數由下式給出：

$$e^{\lambda I + N} = e^\lambda e^N.$$

計算

假設我們想要計算以下矩陣的指數。

$$B = \begin{bmatrix} 21 & 17 & 6 \\ -5 & -1 & -6 \\ 4 & 4 & 16 \end{bmatrix}.$$

它的若爾當型為：

$$J = P^{-1}BP = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 16 & 1 \\ 0 & 0 & 16 \end{bmatrix},$$

其中矩陣*P*由下式給出：

$$P = \begin{bmatrix} -\frac{1}{4} & 2 & \frac{5}{4} \\ \frac{1}{4} & -2 & -\frac{1}{4} \\ 0 & 4 & 0 \end{bmatrix}.$$

我們首先來計算exp(*J*)。我們有：

$$J = J_1(4) \oplus J_2(16)$$

1×1矩陣的指數僅僅是該矩陣的元素的指數，因此$\exp(J_1(4)) = [e^4]$。$J_2(16)$的指數可以用以上提到的公式$\exp(\lambda I+N) = e^{\lambda} \exp(N)$來算出：

$$\exp\left(\begin{bmatrix} 16 & 1 \\ 0 & 16 \end{bmatrix}\right) = e^{16}\exp\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = e^{16}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \cdots\right) = \begin{bmatrix} e^{16} & e^{16} \\ 0 & e^{16} \end{bmatrix}.$$

因此，原矩陣*B*的指數為：

$$\exp(B) = P\exp(J)P^{-1} = P\begin{bmatrix} e^4 & 0 & 0 \\ 0 & e^{16} & e^{16} \\ 0 & 0 & e^{16} \end{bmatrix}P^{-1} = \frac{1}{4}\begin{bmatrix} 13e^{16} - e^4 & 13e^{16} - 5e^4 & 2e^{16} - 2e^4 \\ -9e^{16} + e^4 & -9e^{16} + 5e^4 & -2e^{16} + 2e^4 \\ 16e^{16} & 16e^{16} & 4e^{16} \end{bmatrix}.$$

應用

線性微分方程

矩陣指數在解線性微分方程時十分有用。前面曾提到，以下形式的微分方程

$$\mathbf{y}' = C\mathbf{y}$$

具有解$e^{Ct}y(0)$。如果我們考慮以下向量

$$\mathbf{y}(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$$

我們就可以把線性微分方程表示為：

$$\mathbf{y}'(t) = A\mathbf{y}(t) + \mathbf{b}(t).$$

如果我們作一個猜想，把兩邊乘以一個積分因子 e^{-At}，便得到：

$$e^{-At}\mathbf{y}' - e^{-At}A\mathbf{y} = e^{-At}\mathbf{b}$$
$$\frac{d}{dt}(e^{-At}\mathbf{y}) = e^{-At}\mathbf{b}.$$

如果我們可以計算e^{At}，那麼就得到了微分方程的解。

例子（齊次）

假設我們有以下的微分方程組：

$$\begin{matrix} x' & = & 2x & -y & +z \\ y' & = & & 3y & -1z \\ z' & = & 2x & +y & +3z \end{matrix}$$

相關的矩陣為：

$$M = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix}$$

在以上的例子中，我們計算了矩陣指數

$$e^{tM}=\begin{bmatrix}2e^t-2te^{2t} & -2te^{2t} & 0\\ -2e^t+2(t+1)e^{2t} & 2(t+1)e^{2t} & 0\\ 2te^{2t} & 2te^{2t} & 2e^t\end{bmatrix}$$

因此微分方程組的通解為：

$$\begin{bmatrix}x\\y\\z\end{bmatrix}=C_1\begin{bmatrix}2e^t-2te^{2t}\\ -2e^t+2(t+1)e^{2t}\\ 2te^{2t}\end{bmatrix}+C_2\begin{bmatrix}-2te^{2t}\\ 2(t+1)e^{2t}\\ 2te^{2t}\end{bmatrix}+C_3\begin{bmatrix}0\\0\\2e^t\end{bmatrix}$$

也就是說，

$$\begin{aligned}x&=C_1(2e^t-2te^{2t})+C_2(-2te^{2t})\\ y&=C_1(-2e^t+2(t+1)e^{2t})+C_2(2(t+1)e^{2t})\\ z&=(C_1+C_2)(2te^{2t})+2C_3e^t\end{aligned}$$

非齊次的情況──參數變換

對於非齊次的情況，我們可以用積分因子的方法（類似於參數變換的方法）。我們找到形為$\boldsymbol{y}_{\mathrm{p}}(t)=\exp(tA)\boldsymbol{z}(t)$一個特解：

$$\mathbf{y}_p'=(e^{tA})'\mathbf{z}(t)+e^{tA}\mathbf{z}'(t)$$

$$=Ae^{tA}\mathbf{z}(t)+e^{tA}\mathbf{z}'(t)$$

$$=A\mathbf{y}_p(t)+e^{tA}\mathbf{z}'(t)$$

為了讓$\boldsymbol{y}_{\mathrm{p}}$為方程的解，必須有：

$$e^{tA}\mathbf{z}'(t)=\mathbf{b}(t)$$

$$\mathbf{z}'(t) = (e^{tA})^{-1}\mathbf{b}(t)$$

$$\mathbf{z}(t) = \int_0^t e^{-uA}\mathbf{b}(u)\,du + \mathbf{c}$$

因此，

$$\begin{aligned}\mathbf{y}_p &= e^{tA}\int_0^t e^{-uA}\mathbf{b}(u)\,du + e^{tA}\mathbf{c}\\ &= \int_0^t e^{(t-u)A}\mathbf{b}(u)\,du + e^{tA}\mathbf{c}\end{aligned}$$

其中$\mathbf{c}$由問題的初始條件決定。

例子（非齊次）

假設我們有以下的微分方程組：

$$\begin{matrix} x' & = & 2x & -y & +z & +e^{2t} \\ y' & = & & 3y & -1z & \\ z' & = & 2x & +y & +3z & +e^{2t} \end{matrix}$$

那麼我們有

$$M = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix}$$

以及

$$\mathbf{b} = e^{2t} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

用前面的方法，我們可以得出齊次微分方程的解。由於齊次方程的通解與非齊次方程的特解的和就是非齊次方程的通解，因此我們只需要找到一個特解（用參數變換法）。

我們有：

$$\mathbf{y}_p = e^t \int_0^t e^{(-u)A} \begin{bmatrix} e^{2u} \\ 0 \\ e^{2u} \end{bmatrix} du + e^{tA}\mathbf{c}$$

$$\mathbf{y}_p = e^t \int_0^t \begin{bmatrix} 2e^u - 2ue^{2u} & -2ue^{2u} & 0 \\ -2e^u + 2(u+1)e^{2u} & 2(u+1)e^{2u} & 0 \\ 2ue^{2u} & 2ue^{2u} & 2e^u \end{bmatrix} \begin{bmatrix} e^{2u} \\ 0 \\ e^{2u} \end{bmatrix} du + e^{tA}\mathbf{c}$$

$$\mathbf{y}_p = e^t \int_0^t \begin{bmatrix} e^{2u}(2e^u - 2ue^{2u}) \\ e^{2u}(-2e^u + 2(1+u)e^{2u}) \\ 2e^{3u} + 2ue^{4u} \end{bmatrix} + e^{tA}\mathbf{c}$$

$$\mathbf{y}_p = e^t \begin{bmatrix} -\frac{1}{24}e^{3t}(3e^t(4t-1)-16) \\ \frac{1}{24}e^{3t}(3e^t(4t+4)-16) \\ \frac{1}{24}e^{3t}(3e^t(4t-1)-16) \end{bmatrix} + \begin{bmatrix} 2e^t - 2te^{2t} & -2te^{2t} & 0 \\ -2e^t + 2(t+1)e^{2t} & 2(t+1)e^{2t} & 0 \\ 2te^{2t} & 2te^{2t} & 2e^t \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

進一步簡化，就可以得到原方程的特解。

參見

- 矩陣對數
- 指數函數
- 指數映射
- 向量流
- Golden-Thompson不等式
- 位相型分布

參考文獻

- Horn, Roger A.; Johnson, Charles R., Topics in Matrix Analysis, Cambridge University Press, 1991, ISBN 978-0-521-46713-1.
- Moler, Cleve; Van Loan, Charles F., Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later (http://www.cs.cornell.edu/cv/researchpdf/19ways+.pdf), SIAM Review, 2003, **45** (1): 3–49, ISSN 1095-7200 (https://www.worldcat.org/issn/1095-7200).

外部連結

- MathWorld上 *Matrix Exponential* 的資料，作者：埃里克·韋斯坦因。
- 矩陣指數的教程 (http://math.fullerton.edu/mathews/n2003/MatrixExponentialMod.html)

取自 "http://zh.wikipedia.org/w/index.php?title=矩阵指数&oldid=25499445"

- 本頁面最後修訂於2013年3月13日 (星期三) 05:13。
- 本站的全部文字在創用CC 姓名標示-相同方式分享 3.0 協議之條款下提供，附加條款亦可能應用（請參閱使用條款）。
Wikipedia®和維基百科標誌是維基媒體基金會的註冊商標；維基™是維基媒體基金會的商標。
維基媒體基金會是在美國佛羅里達州登記的501(c)(3)免稅、非營利、慈善機構。