从线性方程谈起

(清华大学数学科学系　章梅荣 教授)

烟台大学与北京大学和清华大学有着密切的联系，我在北大读书时的不少同学、朋友还在烟台大学工作。我应邀来烟台大学作这个演讲，受到了朋友们的热情接待，我感到非常高兴，也深表感谢。

刚才侯仁民主任介绍了，我自己是研究非线性系统的，包括动力系统和非线性微分方程。动力系统是研究**非线性系统**随时间演化的行为的，即是从**动态**的观点来研究系统。今天我的讲演主要是从**静态**的观点来讨论方程的(即方程的解的存在性和结构等等)，并且侧重于对**线性系统**的一些看法，这是研究演化系统的基础。

一、引 子

我们先从最简单的线性方程开始。

例 1　设 $a,b\in R,$　讨论方程$ax=b$。

这是初中一年级，现在可能是小学几年级就要学的最简单的线性方程。因为简单，谁都会解。比如，若$a\neq 0$，把a除去，则$x=b/a$。因此它有唯一的解，并且可以明确地解出来。如果$a=0$会怎样呢？暂时先放一下。

别看这个方程很简单，我们在后面还会不断地回到这个方程上来。在数学上作这样的除法会引发出很多理论，如群论、数域、矩阵，还有要讲到的微分算子等等。它们在形式上五花八门，但道理与这个最简单的方程差别并不太大，也就是说，看上去数学走得很远，但其出发点却是简单的。

再看第二个例子。

例 2　鸭和兔同关在一个笼内，数头有 14 个，数脚有 40 只，问鸭、兔各是几只？

现在我们都知道，这是解线性方程组。我们的祖先在很早时就讨论过很多这种问题和解决这些问题的思想。我小时候是在农村长大的，当地有点文化的大人，也算是“知识分子”或“智者”，在我们很小的时候经常拿这些问题来考我们。如果谁能猜出来或者蒙出来答案来，就会得到夸奖。现在变成了小学二年级学生的奥数题目，在座的每一个同学更会求解了。

这样的题目非常多，这个例子仅是我随手写下的。我们来看看最早怎么求解的，重要的是我们应该从生活经验中把问题的内在数量关系找出来。把鸭子个数画成一个○，把兔子个数画成一个□。因为我们知道一个鸭子有一个脑袋，一个兔子也只有一个脑袋，所以这两个量加起来等于 14：

○+□=14

另外鸭子有两只脚，但兔子有四只脚，这样就得到了另一个关系：

2×○+4×□=40

虽然小学二年级学生还没有学过怎样用消元法去解线性方程组，但他们的知识足以解决这些问题。事实上，只要知道最简单的加法和乘法就够了，然后通过“试验”来解决。先随便瞎蒙，比方说假设鸭子是 1 只（○=1），通过看脑袋，自然兔子就是 13 只（□=13）。现在再来看看脚的个数，这时是

$$2\times 1+4\times 13=54>40$$

这么看脚是太多了。因为每只鸭的脚比每只兔的脚要少，这说明假设的鸭子少了，或反过来讲假设的兔子太多了。于是再来试试比较多一些的鸭子。如果鸭子是 13 只（○=13），那么兔子就只剩下 1 只了（□=1），这时候再算脚才 30 只（$2\times 13+4\times 1=30<40$）。显然是鸭子太多了。由此我们可

以去试试中间的数，通过不断的调整，最终得出鸭子数是 8 只（○=8），兔子数是 6 只（□=6）。

这种解法包含了一种重要的思想，那就是“试验”。这是一切科学（包括数学）的基础。

在描述前面的数量关系时，我们的祖先经常是用画个圆（○），框（□）等来表示，也叫做“天元”、“地数”等。我国国家自然科学基金委员会有一个为数学专设的基金就是用“天元”来命名的，以彰示中国在数学方面的贡献。而西方人用多 x, y 来表示（鸭=x,兔=y），这时例 2 中的数量关系表现为

$$\begin{cases} x+y=14 \\ 2x+4y=40 \end{cases}$$

这是目前通行的二元一次方程组的写法，初中生就要学，而在大学的线性代数中更是要系统研究的。

回到例 2，小孩子能回答这样的问题很是受鼓舞。于是常常倒过来反问村里的“智者”：鸭和兔同关在一个笼内，数头有 14 个，数脚有 41 只，问鸭、兔各是几只？结果是“智者”笑了，你们都知道笑的原因，因为按照这种说法，鸭子 7 只半，而兔是 6 只半。这个例子能说明什么呢？数学不仅仅可以去解决已经提出来的一些问题，也可以反过来检验你的提法、设想是否正确，或者说告诉我们应该如何正确地去思考、提炼问题。

刚才讲到的是在自然数、整数范围里来解二元一次方程组。那么，如何在实系数、复系数范围内求解有更多变元的线性方程组呢？这是我们大学一年级的“线性代数”课程中的基本问题。我们知道有个著名的求解方法，这就是高斯消元法。大学里的学习与中学时期有所不同，不光光学习如何把一个具体问题确确实实地解出来，我们会慢慢地走到更高的高度，更希望从结构性的角度、整体的角度来思考问题并发展理论体系。如对线性方程组，我们会更关注方程组在什么样的情况下会有解，在有解的情况下这些解的结构是什么样等等。

二、Fredholm 原理与谱（特征值）理论

现在我们来谈谈 Fredholm 原理与谱（特征值）理论。

为此先来看一个大学一年级线性代数中的基本问题。

例 3 设 A 是 $m\times n$ 实系数矩阵，$b\in R^m$。找出 $x\in R^n$ 使得

$$Ax=b \tag{1}$$

这是线性方程组，大家都知道标准的结论是什么。但我们要从刚才提到的具体例子中逐步跳出来，要探讨的是这样的方程组什么时候有解？若有解，有多少？解的结构又如何？在实际应用时，怎么去求解？怎么去快速求解？

对于方程组（1），有很多关于这些问题的结论。在我的演讲中，我只讲前面的一些问题，主要是谈什么条件下会有解。这样问题是有意义的，原因是并不是随便给了一个矩阵 A 和一个向量 b，（1）一定会有解 x 存在。为过渡到科学问题中经常涉及到的微分方程，我们把线性代数中的标准结论稍微往前推进一步，看看会发生什么。

这里 R^n、R^m 是标准的欧氏空间，它们有标准的欧氏内积 (x,y)。现在来看如果（1）有解，对 A 和 b 有一些什么必然的要求。设（1）有解 x。则对任意 $a\in R^m$ 有

$$(a,b)=(a,Ax)=(A^Ta,x),$$

这里，A^T 表示 A 的转置矩阵。特别地，如果 a 满足 $A^T a=0$，则 b 必然满足

$$(a,b)=0 \text{ 或 } a\perp b \text{（} a \text{ 与 } b \text{ 正交）。}$$

用空间符号表示即为

$$b\in (Ker(A^T))^{\perp} \text{ (正交补空间)。}$$

这是一个必要性的推导，不是太困难可以证明反过来也对。这说明只有当 b 满足这样的要求时，方程(1)才会有解。但从值域的定义知道，方程（1）有解的条件是 $b\in \mathrm{Im}(A)$。由此我们得到一个非常重要的定理。

定理 1 $\mathrm{Im}(A)=(Ker(A^T))^{\perp}$。

这是在线性方程组里看出来的。在大学数学中继续学下去，我们会多次遇到这样的定理，其中最重要的形式是在线性泛函分析中的线性紧算子的 Fredholm 理论中，以上定理也称为 Fredholm 原理，它可以帮助我们讨论更为复杂的无穷维问题：

例 4 设 X,Y 是 Banach 空间，$A:X\to Y$ 是有界线性算子。讨论

$$Ax=b \tag{2}$$

的可解性问题，其中 $b\in Y$。

这时我们要从有限维过渡到无穷维了。与（有限维的）欧氏空间比较，无穷维空间的最重要区别在于失去局部紧性，而大家可以从多元函数的讨论中体味出紧性的重要性。从某种意义上讲，数学家离开了紧性、完备性等性质，对涉及到“无限”的问题是很难下手的，这是因为现实中大多数问题是通过无限的过程来实现的，而紧性、完备性告诉大家可以通过有限的步骤去逼近无限。由于空间没有局部紧性，这时我们需要对（2）中的算子 A 作些限制。当 A 是所谓的线性紧算子时，类似于定理 1 的 Fredholm 理论在经过适当的解释后同样成立。尽管这些定理看上去形式上与有限维很类似，但其应用面大大地拓宽了，不再仅是一群鸭子、一群兔子的问题，各种微分方程问题，包括常微分方程、偏微分方程、积分方程等等都可以得到应用了。

将 Fredholm 原理应用于例 1 中的简单问题，则当 $a\neq 0$ 时，R 上的线性算子 $Ax=ax$ 是可逆的，$Ker(A^T)$ 是零空间，从而 $(Ker(A^T))^{\perp}=R$，因此对任意 $b\in R$ 方程（1）均有解。即使是 $a=0$ 时的结论也可以这么解释，因为此时 $(Ker(A^T))^{\perp}=\{0\}$ 从而仅当 $b=0$ 时方程 $ax=b$ 才有解。这些与我们最初的分析是吻合的。

我们对线性算子来阐述另一个相关联的观点。在例 4 中，考虑空间 Y 与 X 相同的情形。由于空间的表现形式多种多样，算子 A 的表达形式也多种多样，但我们可以从求解方程的角度用易于理解的“数”来理解、刻画 A。引进一个参数 $\lambda\in C$，考虑类似于（2）的线性方程

$$(A-\lambda I)x=b \tag{3}$$

其中 I 为恒同算子。这时候给一个“定义”：若 λ 使得方程（3）对于任一 b 有唯一的解，则称为 A 的“好值”，反之称为“坏值”。在数学上，它们分别被称为 A 的剩余值和谱。在有限维空间时，谱即是比较好理解的矩阵特征值，在无穷维中谱远远要复杂得多。A 的所有的谱记做 $\sigma(A)$，它是 C 的一个子集。这样对于难于理解的算子 A 赋予了一个容易理解的数集 $\sigma(A)$。按照定义，我们就得到

另外一个定理。

“定理 2” 方程（2）对任意的b有唯一解的充分必要条件是$0 \notin \sigma(A)$。

建立了这样的概念后，可解性问题的难度就被降下来了，因为我们不必要理解整个算子A，而仅需要理解数值性的谱集合$\sigma(A)$。事实上，谱这个概念具有普遍的科学意义，其应用范围超过了想象，例如它不仅仅可以描述求解方程这样的静态问题，而且在描述解的运动性质这样的动态问题中也起着十分重要的作用。

按照这样的观点，例 1 中的事实可以解释为$0 \notin \sigma(A) = \{a\}$，即$a \neq 0$。

现在从大学一年级过渡到二年级，那时会接触到常微分方程。

例 5 设$f(t)$是以2π为周期的连续函数，讨论常微分方程

$$x' + ax = f(t) \qquad （a \in R） \tag{4}$$

的2π-周期解的存在性问题。

这是教科书上一个标准的练习题。结论是：如果$a \neq 0$，则对每个给定的$f(t)$，方程（4）有唯一的以2π为周期的解。标准的做法是：因为方程左边是一个常系数的线性方程，因此可以用简单的常数变易公式把（4）所有的解都写出来，然后再想办法把那个周期解区挑出来。这是解法，现在来谈一谈怎么看待这个结论并与我们已经阐述过的结论作比较。

第一种看法，把它看成无穷维空间中的问题：

$$Z = \{\Phi(t) \mid \Phi\text{是连续可导函数，且}\Phi(t+2\pi) = \Phi(t)\}$$

$$X = \{\Psi(t) \mid \Psi\text{是连续的，且}\Psi(t+2\pi) = \Psi(t)\}$$

显然 $Z \subset X$。引进算子

$$L: Z \subset \mathrm{X} \to X, \quad (Lx)(t) = x'(t),$$

算子L是线性的，在Z, X合适的范数（即距离）下，L成为紧算子。因为Z属于X，谱（特征值）的提法也适用这类算子。对这样的算子，在考虑谱时我们仅讨论其特征值（其原因稍复杂一些）：

$$Lx = \lambda x, x \in Z, x \neq 0。$$

所以

$$x' = \lambda x \tag{5}$$

方程（5）的一般解$x(t) = ce^{\lambda t}$。如要有非零解在Z中，必然有$\lambda = 0$，也就是说算子L的唯一的特征值为$\lambda = 0$。于是当$a \neq 0$时,$-a$就不是L的特征值，从而（4）成为

$$(L - (-a)I)x = f$$

并且有唯一的解$x = (L + aI)^{-1} f$。这是一种看法。

另一种看法：我们也可以用不同的算子。把（4）的左边整体地看成一个算子：

$$M: Z \to X, Mx = x' + ax$$

则当 $a \neq 0$ 时，0 不是 M 的特征值，从而（4）有唯一的解 $x = M^{-1} f$ 。

再看一个二阶常微分方程的例子，它在振动理论中是基本的。

例 6 设 $f(t)$ 是 2π 周期的连续函数，对 $a \in R$，讨论二阶方程

$$x'' + ax = f(t) \tag{5}$$

结论 1：当 $a \neq n^2, n = 0,1,2,\cdots,$ 时，对每一 $f(t)$，（5）有唯一的以 2π 为周期的解。

这结论的运动解释是：对简单弹簧振子的强迫振动，如果频率 a 排除掉一些特殊情形，每一个 2π 周期的外力输入，系统将调制出唯一的 2π 周期振动。这些需排除掉的频率即是所谓的共振频率。

结论 2：当 $a = n^2(n\text{为非负整数})$ 时，（5）有 2π 周期的解的充要条件是 $f(t)$ 满足

$$\int_0^{2\pi} f(t)\cos nt dt = 0 \tag{6}$$

与

$$\int_0^{2\pi} f(t)\sin nt dt = 0 \tag{7}$$

（当 $n = 0$ 时，（7）式是自然成立的。）

对于这些结论，我们仍可以通过常数变易公式把（5）所有的解都写出来从而给出证明。对它们的解释也可以仿照例 5 来进行。结论 1 中是因为 a 不是算子 $Lx = -x''$ 的特征值，从而

$$-(L - aI) = aI - L$$

（表示（5）的左边）是可逆的，所以（5）有唯一的 2π 周期解 $x = (aI - L)^{-1} f$ 。

对于结论 2，$a = n^2$ 为 L 的特征值，$Ker(L - n^2 I)$ 是不难求出的。运用 Fredholm 原理，方程(5)有 2π 解的条件是

$$f \in (Ker(L - n^2 I))^{\perp} = (\{c_1\cos nt + c_2\sin nt \mid c_1, c_2 \in R\})^{\perp}$$

这里的内积要用函数空间中的积分内积，即

$$(f, g) = \int_0^{2\pi} f(t)g(t)dt 。$$

这时就得到积分条件（6）和（7）。

下面再做进一步的拓展。如沿用例 5 中的第二种看法中，也可把 ax 这一项放在算子里。但当 a 不是常系数而是

$$a = q(t) \quad \text{（周期地依赖于时间 } t\text{），}$$

对应的算子为

$$(Kx)(t) = -(x''(t) + q(t)x(t))$$

此时需要考虑

$$Kx = \lambda x$$

即

$$x'' + (\lambda + q(t))x = 0 \tag{8}$$

当 $\lambda = 0$ 时，（8）成为

$$x'' + q(t)x = 0 \tag{9}$$

这时的方程(9)称为 Hill 方程。由于（8）和（9）不能直接积出所有的解，这时 K 的谱（特征值）问题（8）就不那么容易简单讨论清楚了。一个经典而又重要的理论是当 $q(t)$ 是周期函数时，（8）的谱问题是清楚的。但是当（势函数）$q(t)$ 变得复杂一些，如几乎周期势函数，谱问题（8）成为数学物理中迄今为止远没有得到充分理解的大问题。因此即使是线性问题也远非简单。

由此，Fredholm 原理和谱（特征值）理论在解的存在性问题中的巨大作用可见一斑。

三、非线性问题

最后，我们谈一点非线性方程问题。诚如大家所知，自然界中更多的是非线性问题，包括静态的和动态的。一般来讲，好的线性理论往往能给出对应的线性问题的比较完整的解答。但对非线性问题，大多数是没有完整的回答的。因此在解决非线性问题时要运用多种理论和方法，也需要大量的线性问题的工作基础。

例 7 （局部问题）考虑 $f : R^n \to R^n$ 是非线性连续函数（也可把 f 看为一个系统）。假设

$$f(x) = 0 \tag{10}$$

有解，比如 $x = x_0 = 0$ 。讨论当一个输入 $b \in R^n$ 比较小时，方程

$$f(x) = b \tag{11}$$

是否有解？

在微积分中，我们有一个结果，即

反函数定理 当 f 是连续可微分且 $A = D_x f(0)$ 是可逆的（即 $0 \notin \sigma(A)$ ）时，则当 b 比较小时，方程（11）在 $x = 0$ 附近有唯一的解。

这个定理中的核心条件是“ A 是可逆的”，它说明线性问题（2）对任意 b 的可解性。但同时与（2）的消元法比较，反函数定理却无法给出简单可操作的求解方法。另外，如果失去核心条件，（11）成为所谓的分支问题，就会变得更加复杂，使得象线性代数里完整解决问题(2)的设想对（11）几乎无法实现。

我们再来看看**整体**的（也叫**大范围**的）非线性问题。这样的问题的重要性是不言而喻的，但其解答是十分困难的同时也是现代数学所致力于研究的。原则上，当对应的线性问题是不退化的（与谱理论有关），则线性的结果在适当小的非线性扰动后仍部分保留下来。因此当我们对对应的线性问题有充分的认识后，总能对非线性问题给出一部分解答。但如何理解“适当小的非线性扰动”，总是有很多不同看法。

例 8 设$a \neq 0$，则已知方程（1）对任意b有唯一解（（1）的非退化性）。设$g{:}R{\to}R$是连续函数，讨论非线性方程

$$ax + g(x) = b \tag{12}$$

是否有解。

结论 1：由于我们知道（12）所对应的线性问题（1）是非退化的，因此只要$g(x)$是适当小时，则（12）会有解。例如，当$g(x)$在R上是有界的，则（12）至少有一个解。这很容易用微积分来证明，如设$a > 0$，则在x趋于$\pm\infty$时，函数$ax + g(x)$的值也相应地趋于$\pm\infty$，从而从介值定理即知（12）会有解。但与（1）不同，此时（12） 的解的唯一性一般是不能保证的。

结论 2：对“适当小”有不同的理解。例如，$g(x)$在R上也可以是无界的，但又满足

$$\sup_{x \in R} | g'(x) | < | a | \tag{13}$$

则（12）会有解。条件（13）也可以用$g(x)$的李普希兹条件来做要求，……

在本例中我们可以看到，$a \neq 0$的这个线性问题的条件是非常重要。如$a = 0$上，（12）又回到了复杂的（11）了。

如果我们对线性问题的理解有足够的自信，如基于例 6 中的结论 1 我们完全可以对如下的微分方程问题推断出“大胆”的猜测。

例 9 设$a \neq n^2, n = 0,1,2,\cdots$，$g{:}R{\to}R$是连续有界函数，则对每个$2\pi$周期的连续函数$f(t)$，方程

$$x'' + ax + g(x) = f(t) \tag{14}$$

至少有一个2π周期解。

这个猜测确实是对的，其证明远比例 8 中的结论 1 要困难了，但仍是可以用多种方法来实现的，如微分方程的相平面分析方法，非线性泛函分析中的拓扑度方法、Leray-Schauder 度方法、变分方法等。仿照例 8 中的结论 2，我们也可以推断出只要$g(x)$满足

$$\sup_{x \in R} | g'(x) | < M \tag{15}$$

且M比较小时，则（14）会有2π周期解。至于M到底要多小，会与（5）中左边的线性算子的谱有密切关系。同样地，条件（15）也可以改用$g(x)$的李普希兹条件。

四、结 论

今天我讲到了线性与非线性之间的联系，也提到了一些静态和动态、局部与整体之间的关系，但主要集中在第一个方面。我的体会是，当你需要从动态的观点去研究整体的非线性问题时，你必须对从静态的观点去研究线性问题有足够的理解。从其它角度来研究非线性系统，其道理也该是如此。看似简单的线性理论正是我们迈向非线性问题的基础，当你越多地了解了线性的基础，你就越

有可能达到非线性的高峰。因此，要重视“简单”的线性理论。

今天就到这儿。感谢烟台大学的邀请，谢谢大家！

致谢：作者对烟台大学张志军教授为整理此演讲稿所作的贡献深表谢忱！