$$y=\left(\cos\frac{3}{2}x\right)^{\frac{1}{3}}\cdot\left[c_1e^{\frac{\sqrt{3}}{2}x}+c_2e^{-\frac{\sqrt{3}}{2}x}\right]$$

为通解.

例 3. 解方程 $(2x+a)y''+y'+by=0$, 其中 a、b 为常数.

解: 满足 VI, 所以

$$y=\begin{cases}c_1\cos\sqrt{b(2x+a)}+c_2\sin\sqrt{b(2x+a)}, & b>0\\ c_1e^{\sqrt{-b(2x+a)}}+c_1e^{-\sqrt{-b(2x+a)}} & b<0\end{cases}$$

为通解.

例 4. 解方程 $y''+(e^x-1)y'+e^{2x}y=0$

解. 满足 X, 所以

$$y=e^{\frac{1}{2}(1-e^x)}\left[c_1\cos\frac{\sqrt{3}}{2}(e^x-1)+c_2\sin\frac{\sqrt{3}}{2}(e^x-1)\right]$$

为通解

例 5. 解方程 $y''-\left(x+\frac{1}{x}\right)y'+\frac{x^2}{4}y=0$

解. 满足 XI, 所以

$$y=e^{\frac{x^2}{4}}\left[c_1+c_2\frac{x^2}{4}\right]$$

为通解.

例 6. 解方程 $y''+\frac{1+e^{2x}}{1-e^{2x}}y'+\frac{(1-e^{2x})^2}{(1+e^{2x})^2}y=0$

解: 满足 IX, 所以

$$y=(e^x+e^{-x})^{\frac{1}{2}}\left\{c_1\cos\frac{\sqrt{3}}{2}[x-\ln(1+e^{2x})]+c_2\sin\frac{\sqrt{3}}{2}[x-\ln(1+e^{2x})]\right\}$$

为通解.

例 7. 解方程 $y''+(e^{\frac{x}{2}}-1)y'+\left(1+\frac{1}{4}e^x-\frac{1}{4}e^{\frac{x}{2}}\right)y=0$

解. 满足 XIV, 所以

$$y=e^{1+\frac{1}{2}x-e^{\frac{x}{2}}}\left[c_1\cos\frac{\sqrt{3}}{2}x+c_2\sin\frac{\sqrt{3}}{2}x\right]$$

为通解.

例 8. 解方程 $y''+\frac{4}{x^2}y'+\left(\frac{4}{x^4}-\frac{4}{x^3}-4\right)y=0$

解. 满足 XVI, 所以

$$y=e^{\frac{2}{x}}[c_1e^{2x}+c_2e^{-2x}]$$

为通解.

例 9. 解方程 $y''+\frac{1}{e^{\frac{1}{2}x}-1}y'-\frac{1}{4(e^{\frac{1}{2}x}-1)}y=0$

解: 满足 XVII, 所以

$$y=\frac{e^{\frac{1}{2}x}}{e^{\frac{1}{2}x}-1}(c_1+c_2x)$$

为通解.

例 10. 解方程 $y''+\frac{1}{x}y'+\left(1-\frac{1}{4x^2}\right)y=0$

解: 满足 XIX. 所以

$$y=x^{-\frac{1}{2}}(c_1\cos x+c_2\sin x)$$

为通解.

参考文献

[1] 贺建勋. 常微分方程中册, 湖南科技出版社. 1980 年 P393—394.

[2]. 钱详征. 常微分方程解题方法. 湖南科技出版社 1984 年 P243.

关于一个优选问题的思考*

陈木法

(北京师范大学数学系)

无论在工作或日常生活中, 总是希望把事情做好, 即采用最好的方案. 如果对于所从事的工作已有透彻的了解, 这比较容易办到. 用数学的语言来说, 如果所关心的函数 $f(x)$ 是已知的, 那么可以求出它的最大值点. 再具体些, 如果已知 $f(x)=-x^2$, 那么这个函数在 $x=0$ 处取最大值 0. 答案十分简单. 然而, 在实际中, 情况远非如此. 就拿煮米饭为例, 水放多了会把干饭煮成稀饭, 放少了煮不熟, 没法吃. 因而"煮干饭水放多少最好吃"就是一个优选问题.

* 本文是为我系90届新生举办讲座的讲稿. 目的是以一个数学问题为例, 介绍笔者学数学、用数学的点滴体会

在这里，“好吃”这个函数是什么？它的答案至少是因人而异的．在实践中，总是通过试验，找出最好的放水量．差别在于，有些人较快就能学会，而另些人则要花较长时间．这就提出问题，如何以最快的速度(即最少的试验次数)找出最佳方案？华罗庚把这类数学方法称为优选法；不仅要选优，而且要以最优方法去选优．

我们仅限于讨论单因素的优选问题，并假定试验初始区间是$[0,1]$．希望以最快的速度，找出$[0,1]$上的点x_0，使得$f(x_0)$为$f(x)$在$[0,1]$上的最大值．由于未知函数$f(x)$的种类太多，需要加以限制．这也是把实际问题转化为数学问题的第一步，抽出本质的部分，排除非本质部分．我们将限于一类“好”的函数，单峰函数．

定义．区间$[0,1]$上的函数$f(x)$称为单峰函数，如果在$[0,1]$上有唯一的一点c_f，使得

$$f(x)\leqslant f(c_f),\quad \forall x\in[0,1];$$

而且$f(x)$在$[0,c_f]$和$[c_f,1]$上严格单调．此时，称c_f为$f(x)$的峰值点．

用记号$\mathscr{F}$表示$[0,1]$上单峰函数的全体．

为什么说单峰函数“好”？因为它有两个特点：一是它包含唯一的峰值点，这是我们所关心的目标；二是每当做完两次试验，总可以通过比较消去一部分区间(见图1)．使我们能够一步一步地逼近峰值点．为方便起见，把每步消去一段之后所剩下的区间称为剩余区间．

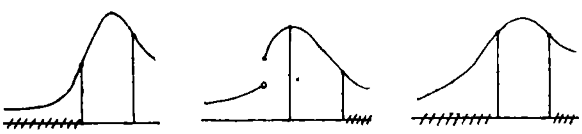


图1

其次，关于试验方法也需要明确．

定义．所谓试验策略$\mathscr{D}$，仍是这样一个规则：在第一步，它确定一个试验点$x_1=x_1(\mathscr{D})$，不依赖于具体的$f\in\mathscr{F}$；在第n步，它根据前$n-1$步的试验点$x_1,\cdots,x_{n-1}$及试验结果$f(x_1),\cdots,f(x_{n-1})$确定第n个试(验)点$x_n=x_n(\mathscr{D};x_1,\cdots,x_n;f(x_1),\cdots,f(x_{n-1}))$．

在进一步讨论之前，先给出试验策略的例子．为此，先给出一串分数

$$\frac{1}{2},\frac{2}{3},\frac{3}{5},\frac{5}{8},\frac{8}{13},\cdots\cdots.$$

其规则是

$$F_0=F_1=1,F_n=F_{n-1}+F_{n-2},n\geqslant 2.$$

而第n个分数为F_n/F_{n+1}．

定义．给定$n\geqslant 1$，令

$$x_1=x_1(\mathscr{F}_n)=F_n/F_{n+1}.$$

从第2步开始，依照对称规则安排新试点．例如x_2是x_1关于中点的对称点(见图2左)

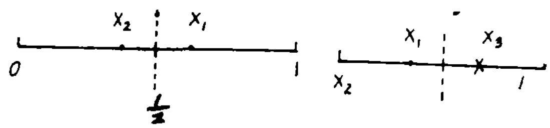


图2

假定两次试验后的剩余区间是$[x_2,1]$那么，取x_3为x_1关于中点的对称点(见图2右)．依此类推．我们称这种试验策略为Fibonacci搜索法或分数法．

最后，还需要引进一个标准，以比较两种不同的试验策略的优劣．设$\mathscr{D}$是一个n步试验策略，即使用策略$\mathscr{D}$，恰好做n次试验．在这n次试验中，有n个试点$x_1,x_2,\cdots,x_n$及n个试验结果$f(x_1),\cdots,f(x_n)$．其中必有一个试点，记作$c_f(\mathscr{D},n)$，使f达到最大，即

$$f(c_f(\mathscr{D},n))\geqslant f(x_i),i=1,2,\cdots,n.$$

但由定义，f的峰值点是c_f．一般地讲，$c_f(\mathscr{D},n)$与c_f有一定的距离．命

$$\delta(\mathscr{D},n)=\sup_{f\in\mathscr{F}}|c_f-c_f(\mathscr{D},n)|.$$

定义．称$\delta(\mathscr{D},n)$为试验策略的n步精度，并称试验策略$\mathscr{Q}$为n步最优的．如对任意的n步试验策略$\mathscr{D}$，

$$\delta(\mathscr{D},n)\geqslant\delta(\mathscr{Q},n).$$

作为练习，请读者自行算出分数法的各步精度，加深对于以上概念的理解．于此，可对上述概念作些注释．对于给定的$\mathscr{D}$，当f不同时，n步所得到的最好点$c_f(\mathscr{D},n)$与真正的峰值点c_f的距离是不同的．换言之，$|c_f(\mathscr{Q},n)-c_f|$随f的变动而变动．最坏的情形是使$|c_f(\mathscr{Q},n)-c_f|$达到“最大”的那种f．换言之，在最坏的情况下，两者相距

$$\sup_{f\in\mathscr{F}}|c_f(\mathscr{D},n)-c_f|.$$

然后，要从最坏情形中找出最好的试验策略，即找出$\mathscr{Q}$，使得

$$\delta(\mathscr{Q},n)=\inf_{\mathscr{D}}\sup_{f\in\mathscr{F}}|c_f(\mathscr{D},n)-c_f|.$$

直观上讲,从最坏的可能性中找出最好的结果,这样做最保险. 这种思想称为"极大极小化"原理. 在近代数学中, 这一原理占有重要地位.

有了以上准备, 我们可以介绍第一个著名的结果.

定理((J. Kiefer[1], 1953). 分数法 $\mathscr{F}_n$ 是 n 步最优策略.

Kiefer 在他的上述短文的结尾, 已经指出分数法的一个缺点, 需要事先确定好试验的次数,以后简称为"预定次数". 否则,在 n 步试验之后,

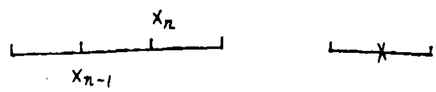


图3

"好点" 位于剩余区间的中点(见图3). 如想再做下去,已无法依同一对称规则安排. 为克服这个缺点,他建议以

$$\omega = \lim_{n\to\infty}\frac{F_n}{F_{n+1}} = \frac{\sqrt{5}-1}{2} \doteq 0.618,$$

作为第 1 个试点; 然后依照对称规则安排以后试点. 因为 ω 是黄金分割常数,这个试验策略称为黄金分割法, 记作 $\mathscr{W}$. 华罗庚把它简称为 0.618 法. 这个方法在我国是相当普及的, 在生产实践中有着广泛的应用.

总之,Kiefer 最先找到 n 步最优策略 $\mathscr{F}_n$, 把 $\mathscr{W}$ 视为 $\mathscr{F}_n$ 的近似方法. 到了六十年代,华罗庚提出不同看法. 他认为"黄金分割法"$\mathscr{W}$是最优的,而"分数法"是一种特殊的方法,它是前者的一种近似. 先谈谈后一种看法, 把$\mathscr{W}$展开成连分数

$$\omega = \cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\ddots}}}.$$

如果取前 $n+1$ 个连分数作为 ω 的近似值,则得出分数

$$F_n/F_{n+1} = x_1(\mathscr{F}_n).$$

它正是分数法的第一个试点. 用分数列 $\{F_n/F_{n+1}\}$ 来逼近无理数 ω,逼近速度较快. 这用到数论中的丢番图逼近.

当然,华罗庚的最优性与 Kiefer 的意义不同. 先引入

定义. 称 $\mathscr{D}$ 为对称策略, 如从第 2 步开始,它依对称规则安排新试点.

华的结果是

定理 (华罗庚[2,3]). 对于任何对称策略 $\mathscr{D}$,当 n 充分大时,总有

$$\delta(\mathscr{D},n)\geqslant\delta(\mathscr{W},n).$$

换言之,黄金分割法是无穷远处的最优策略.

上述结果对于任意策略也是对的, 详见洪加威[4].

让我们再回过头来分析上述两种不同的看法. 首先, 想大家会赞同排除"预定次数"这一限制条件. 即应当寻求这样一种试验策略, 它不仅对于任何单峰函数而且对于任何的试验次数而言都是最优的. 其次,华的想法是考虑无穷的最优性, 因而不必再顾虑究竟会做多少次试验. 从这个角度来看,两者的区别在于,Kiefer 把试验次数预定于有限数,而华则预定于无穷. 因此,两者都是"预定次数"的.

现在, 想想如何考虑适用于任何试验次数的最优策略. 这首先需要引入新的精度概念. 按照前面所述的"极大极小化"原理, 应寻找这样的策略,使之达到

$$\inf_{\mathscr{D}}\ \sup_{n\geqslant1}\ \sup_{f\in\mathscr{F}}\ |c_f - c_f(\mathscr{D},n)| = \inf_{\mathscr{D}}\ \sup_{n\geqslant1}\ \delta(\mathscr{D},n).$$

这种想法十分自然但却不合理. 原因在于, 对于不同的 n,试验区间的尺度不是同一量级的,因而 $\delta(\mathscr{D},n)$ 也不属于同一量级. 较合理的做法是代之以相对误差. 我们知道, n 步最优策略 $\mathscr{F}_n$ 有最佳精度.

$$\delta(\mathscr{F}_n,n) = 1/F_{n+1}.$$

对于任一策略 $\mathscr{D}$, 不可能指望 $\mathscr{D}$ 在每一步都达到最佳精度. 然而,对于每一个 n,总可以比较 $\mathscr{D}$ 与 $\mathscr{F}_n$ 精度的相对误差

$$\frac{\delta(\mathscr{D},n)-\delta(\mathscr{F}_n,n)}{\delta(\mathscr{F}_n,n)} = F_{n+1}\delta(\mathscr{D},n)-1.$$

这引导我们使用

$$\delta(\mathscr{D}) = \sup_{n\geqslant1} F_{n+1}\delta(\mathscr{D},n)$$

作为策略$\mathscr{D}$的精度. 称试验策略 $\mathscr{Q}$ 是最优的, 如果对于任一试验策略 $\mathscr{D}$,有 $\delta(\mathscr{D})\geqslant\delta(\mathscr{Q})$.

在这种一般的精度意义下,有

定理([6], 1977). 对于任何对称策略 $\mathscr{D}$, 有

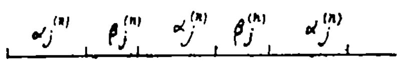
$$\delta(\mathscr{D})\geqslant\delta(\mathscr{W}).$$

换言之，黄金分割法是在不定次数意义下的最优策略.

上述结果只是表明，$\mathscr{W}$ 在这种更为广泛的意义下也是最优的. 这只不过是对 $\mathscr{W}$ 的最优性作出的一种新的解释. 如果只能做到这一点，那么这项工作也就没有太多价值. 但事情并非如此完结.

现在假定有偶数（例如说 $k=2i$, $i\geqslant1$）部机器. 因此每步可安排 $2i$ 个试点. 首先，固定试验的步(批)数 n.

定义 给定 $i\geqslant1$ 和 $n\geqslant1$. 定义每批 $2i$ 个试点的 n 步试验策略 $\mathscr{E}_n^{(i)}$ 如次：在第 1 步，将 [0,1] 区间分成长、短相间的诸段

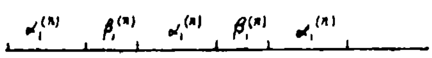


此处

$$\alpha_1^{(n)}=\frac{2(i+1)^{n-1}-1}{2(i+1)^n-1},$$

$$\beta_1^{(n)}=\frac{1}{2(i+1)^n-1}.$$

而在以后各步，使用同一构造



虽然 $\alpha_j^{(n)}$ 与 $\beta_j^{(n)}$ $(j\geqslant2)$ 与前面的 $\alpha_1^{(n)}$ 和 $\beta_1^{(n)}$ 不同.

限于对称策略的原因有二：一是所有的最优策略都是对称的；二是构造复杂的策略在实践中没有多少用处. 与前面类似，我们可以定义一个试验策略的精度等概念. 可以证明，$\mathscr{E}_n^{(i)}$ 是 n 步最优策略. 这类似于每步一个试点的分数法 $\mathscr{F}_n$.

回忆 $\mathscr{W}$ 是作为 $\mathscr{F}_n$ 的极限得出的：

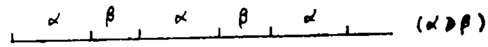
$$w=\lim_{n\to\infty}x_1(\mathscr{F}_n)=\lim_{n\to\infty}F_n/F_{n+1}.$$

但在目前的情况下，

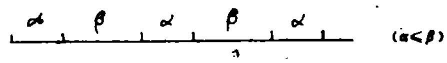
$$\lim_{n\to\infty}\alpha_1^{(n)}=\lim_{n\to\infty}(\alpha_1^{(n)}+\beta_1^{(n)})=\frac{1}{i+1}.$$

因此，此法失效. 事实上，对于每批偶数个试点的情况，不存在无穷远处的最优策略. 详见洪加威([5],1974).

定义 称试验策略 $\mathscr{D}$ 为基本的，如果在每一步，都有如下结构



或者



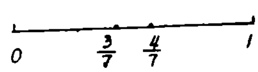
特别地，如每批为 $2i$ 个试点且在每一批有 $\alpha=\beta$，则记之为 $\mathscr{E}^{(i)}$.

定理（[6],1977）假定每批有 $2i$ $(i\geqslant2)$ 个试点，则对于任何基本策略 $\mathscr{D}$，有

$$\delta(\mathscr{D})\geqslant\delta(\mathscr{E}^{(i)}).$$

换言之，$\mathscr{E}^{(i)}$ 是每批 $2i$ 个试点的最优策略.

这个结果第一次给出 $\mathscr{E}^{(i)}$ 的最优性. 然而，$\mathscr{E}^{(i)}$ 的构造简单. 每批将试验区间等分，而在等分点上安排试验. 似乎不难猜到. 但我们着重指出，对于 $i=1$ 的情形，即每批两个试点，第 1 批试点应改取为



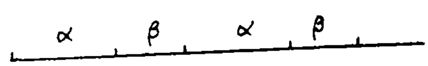
随后各批依然使用等分规则，才是最优的[6].

最后，每批 $2i-1$ $(i\geqslant1)$ 个试点的情况. 如同 $i=1$ 的情形，可引进序列

$$F_0^{(i)}=F_1^{(i)}=1,$$

$$F_n^{(i)}=i(F_{n-1}^{(i)}+F_{n-2}^{(i)}),\ n\geqslant2.$$

然后定义基本策略 $\mathscr{F}_n^{(i)}$ 如次，在第 1 步(批)，



$\alpha=F_n^{(i)}/F_{n+1}^{(i)}$, $\beta=1/i-\alpha$. 以下各步依然分为长、短相间的 $2i+1$ 段(在 $2i$ 个分点上有一已试点). 可以证明，$\mathscr{F}_n^{(i)}$ 是 n 步最优策略. 再一次回到早先导出 $\mathscr{W}$ 的手法，命

$$\omega(i)=\lim_{n\to\infty}F_n^{(i)}/F_{n+1}^{(i)}=\frac{\sqrt{i(i+4)}-i}{2i}.$$

我们可得出类似于 $\mathscr{F}_n^{(i)}$ 的试验策略 $\mathscr{W}^{(i)}$，它是无穷远处的最优策略[5].

（下转44页）

以前在讲这个问题的前一堂课上，我让每个同学写一个纸条，写上自己的姓名和出生年月．在下次讲古典概型时，向同学们提出一个问题：某房间有 n 个人，他们中至少有两个人在今年的同一天庆祝生日的概率是多少（2 月 29 日生的人除外）？这一问题立即引起同学们的兴趣，经过分析，知道该事件的概率为

$$P=1-\frac{A_{365}^{n}}{365^{n}}$$

接着公布在他们班统计的结果．他们往往认为出现两个人同一天过生日是偶合，但进一步计算得知，当 $n>23$时，$P>\frac{1}{2}$；$n>50$时，$P>0.97$；$n>100$ 时，$P=1-\frac{1}{3\times10^{5}}\approx1$，几乎成了必然事件．每次统计的结果，在一个班的同学中几乎都有两个以上的人在同一天过生日，这个结果加深了同学们对问题的理解，同时，有利于他们对具体问题具体分析，攻克难点，掌握古典概型的计算．

随机变量函数（以连续型为例）的分布是概率论的另一难点．它是在已知作为自变量的随机变量 x 的密度函数 $f(x)$（其反函数 $x=h(y)$ 在对应区间上存在，且单调可导）的情况下，求作为函数的随机变量 $y=g(x)$ 的密度函数 $\psi(y)$．

要讲清这个问题必须交代两点：

1）密度函数和概率无直接关系，只有转化为相应的分布函数，才能与概率联系起来．

2）事件相等，概率必相等．

通过以下关系

$$F_y(y)=P\{y\leqslant y\}=P\{g(x)\leqslant y\}=P\{x\leqslant h(y)\}$$

分布函数概念换元事件相等概率相等

$$=\int_{-\infty}^{h(y)}f(x)dx \quad (*)（假定 f(x) 单增）$$

分布函数与密度函数的关系

由变上限复合函数求导法则，由(*)式两边对 y 求导得

$$\phi(y)=F'_y(y)=f[h(y)]h'(y)$$

据笔者体会：

1）只要讲清以上四个概念，被认为异常抽象的概率概念部分的问题就能得到很好的解决．

2）用现代方法论观点对教材进行科学的分析，可以使我们分清问题的重点、难点，在制订教学计划时可以做到心中有数，在教学过程中可以强化相应部分的讲解工作，以便学生深刻地领会该门学科．

3）概率论是一门实用性很强的科学，在讲授这门课时，选择生动的实例可以使学生把一些重要的概念铭记在头脑中．

参考文献

1. 《数学方法论选讲》（大连理工大学徐利治教授编写）
2. 《概率论与数理统计》（浙江大学数学系编）

（上接38页）

问题是：$\mathscr{W}^{(i)}$ 是否也是不定步（批）数意义下的最优策略？这个问题最近才获得解决．答案是否定的．这种情形下的最优策略已经找到，可惜不能在此讨论．

参 考 文 献

[1] *Kiefer J., Sequential minimax search for a maximum, Proc. Amer. Math. Soc.* 4 (1953), 502—506.
[2] 华罗庚，优选法平话及其补充，国防工业出版社，1970.
[3] 华罗庚，优选学，科学出版社，1980.
[4] 洪加威，论黄金分割法的最优性，数学的实践与认识，2 (1973).
[5] 洪加威，论批数不限定情况下一维优选问题的最优策略，中国科学，2(1974).
[6] 陈木法，论不定次（批）数条件下单因素优选问题的最优策略，贵阳师院（今贵州师大）学报，3(1977)，117—134.