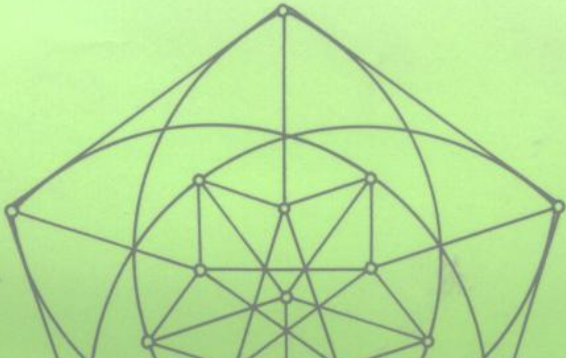
数学模型基础

王树禾　编著



中国科学技术大学出版社

数学模型基础

王树禾　编著

中国科学技术大学出版社
1996·合肥

内 容 简 介

本书选择众多实际问题，建立它们的数学模型，且给出解答．这些问题来自科学技术、社会生活、军事战争、交通通讯、医疗卫生、计划生育、生态环保、天气预报、对策运筹、试验设计、密码编译、选举评优、投资发展、广告促销、物价问题、经济调整等等领域，或十分重要，或十分敏感．通过这批非人工的“真问题”之研究，使读者强化化实际问题为数学模型的意识和建模解题的能力与技巧．

本书可做为理工农医财各专业大学生与研究生选修课教材，也供数学模型竞赛选手及有关的科技工作者参考．

数 [illegible]

王树禾 编著

中国科学技术大学出版社出版发行

（安徽省合肥市金寨路 96 号，230026）

中国科学技术大学印刷厂印刷

全国新华书店经销

*

开本：850×1168／32　印张：10.125 字数：260 千

1996 年 5 月第 1 版　　1996 年 5 月第 1 次印刷

印数：1—6000 册

ISBN 7-312-00746-5／O·17　　定价：8.00 元

前　言

当今社会日趋数学化，随着人类生活质量的提高，生产力的发展和科学文化的进步，数学迅速介入一切领域，高技术实则为一种数学技术.所谓数学技术，即指把现实的问题转述成一个相应的数学问题，且用计算机加以解决或用数学理论定性定量加以研究，得出那个现实问题的定量结论或重要性质.

但是，如何建立和解答纷纭复杂的实际问题的数学模型，我们很难给出普遍的原则，本书以众多典型的具体问题之建模分析，向读者显示建立数学模型的思路与技巧.所谓数学模型，乃是现实世界当中某一类特殊的运动变化过程或关系或结构的一种模拟性的数学结构，是对现实模型的理想化，是一种科学的抽象过程.

在建模过程当中，不得不把那些对我们关心的问题影响甚微的因素忽略掉，不然，所得模型因为数学结构太复杂而失去数学上的可解性；但也不能把足够相关的因素忽略掉，不然所得模型因为不能足够准确地反映实际情况而失去可靠性.可解性与可靠性同时最佳是很罕见的，一般我们总是在可解性的前提下，力争有满意的可靠性.

同一个现实问题，可以建立不同的几个数学模型，事实上，被我们研究的实际问题好比一个“黑匣子”，要用各种方法，以各种角度去观察探讨它的秘密，有时，建模(的思考方式)酷似瞎子摸象式的进行.数学模型的建立需要有创造性、想象力甚至具有一定的艺术性，必须接受实践的检验，有时需要反复修正.

本书涉及的知识面极宽，在写作过程中，我们一般不再交代所用的数学知识，采取拿过来就用的态度，这些知识主要来自线性代数、微积分、解析几何、图论、运筹学、组合数学、常微分方程、偏微分方程、数理统计等等，我们假设读者对这些数学分支已经通晓，如果有

必要，您可以选取本书列出的参考文献去查阅. 知识是重要的，但能力(主要是想象与抽象能力)更为重要.

本书是大学高年级或硕士研究生选修课的教科书，也可供各行业科技人员参考，如果做教材，周学时 4，一学期可授完，亦可做大学生建模竞赛集训班的参考书.

作者的知识、能力和实践经验有限，书中疏漏错误恐怕不少，敬候读者指正.

王树禾

1995 年 7 月

目　次

第一篇　离散数学模型

1.1　最短路线问题

给定连接若干城市的铁路网，找一条给定两城市间的最短路线，此即所谓**最短路线问题**，它的数学模型如下：

$G(V,E)$是加权图，即$\forall e\in E(G)$，$\exists\, w(e)\in \boldsymbol{R}^+$，记

$$W(G)=\sum_{e\in E(G)} w(e),$$

$\forall$ u,v$\in$V(G)，求轨 P_0(u,v)，使得

$$W(P_0)=\min_{P\in \mathscr{P}(u,v)}\{W(P)\}$$

其中$\mathscr{P}(u,v)$是从 u 到 v 的轨集合.

以后我们称上述 $P_0(u,v)$之长 $W(P_0)$为顶点 u,v 之**距离**.

边权在友谊图中代表两位朋友感情深厚的程度；在通讯图中，边权表示通讯线路的造价或维修费用，等等.许多不同的实际问题，其数学模型与上述铁路网上最短路线问题的数学模型一致.

为解决最短路线问题，下面介绍顶点 $u_0\in V(G)$到连通图 $G(V,E)$各顶点最短轨的一种有效算法.所谓算法，指一组有穷规划，它准确告知，为解决给定的问题，何时应做何种操作.

一个图论算法的计算量为 $O(P(\nu,\varepsilon))$时，称其为有效算法或好算法，其中 $P(\nu,\varepsilon)$是多项式 ν 与 ε 分别是图的顶数与边数.

Dijkstra 算法：(u,v 不相邻时，$w(uv)=\infty$)

(1)　令 $l(u_0)=0, l(v)=\infty, v\neq u_0, S_0=\{u_0\}, i=0$.

(2)　对每个 $v\in \overline{S_i}$ ($\overline{S_i}=V(G)-S_i$)，$\min\{l(v), l(u_i)+w(u_iv)\}$代替 $l(v)$；设 u_{i+1}是使 $l(v)$取最小值的$\overline{S_i}$中的顶，令 S_{i+1}

$=S_i \cup \{u_{i+1}\}$.

(3) 若 $i=\nu-1$,止;若 $i<\nu-1$,用 $i+1$ 代替 i,转(2).

由上述算法知:

(1) S_i 中各顶标 $l(u)$ 即为 u_0 到 u 的距离.又因 $\nu<+\infty$,故有限步之后,$V(G)$中每一顶都标志了与 u_0 的距离,从而可以找到各顶到 u_0 的最短轨.

(2) Dijkstra 算法的时间复杂度为 $O(\nu^2)$,所以是有效算法.

例如图 1 中由顶 u_0 到各顶的最短轨道及距离(各顶标)已标志在(h)中,u_0 至各顶的最短轨可按图中粗实线找到.

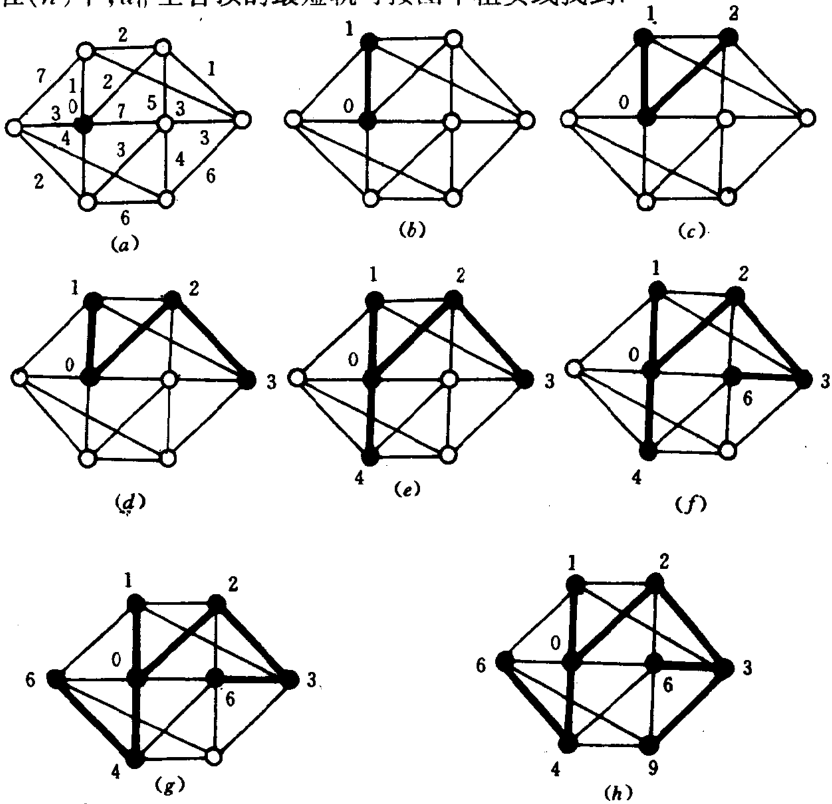


图 1

1.2 连接问题

今欲修筑连接几个城市的铁路,已知 i 城与 j 城之间铁路造价为 c_{ij},试设计一个路线图,使总造价最低.

上述连接问题的数学模型是:在已知的加权连通图上求最小权的生成连通子图,它显然是一个生成树,这个生成树也叫最轻树或最优树,下面介绍求最优树的算法.

Kruskal 算法:

(1) 选 $e_1 \in E(G)$,使得 $w(e_1)=\min$.

(2) 若 $e_1 e_2,\cdots,e_i$ 已选好,则从 $E(G)-\{e_1,e_2,\cdots,e_i\}$ 中选取 e_{i+1},使得

(i) $G[\{e_1,e_2,\cdots,e_i,e_{i+1}\}]$中无圈,且

(ii) $w(e_{i+1})=\min$.

(3) 继续进行到选出 $e_{\nu-1}$为止.

其中 $G[E'](E'\subseteq E(G))$叫做 E'的导出子图,它是以 E'为边集,以 E'中边之端点为顶的图.

定理 Kruskal 算法选得的边的导出子图为最优树.

证 Kruskal 算法得出的子图 T^* 是生成树自不待说,下证它的最优性.设 T^* 不是最优树,T_1 是 G 的任给定的一个生成树,$f(T_1)$是$\{e_1,e_2,\cdots,e_{\nu-1}\}$中不在 T_1 上的 e_i 的足标 i 的最小值,令 T 是使 $f(T)$最大的一个最优树,因为 T^* 不是最优树,又 $E(T^*)=\{e_1,e_2,\cdots,e_{\nu-1}\}$,故 $e_1,e_2,\cdots,e_{\nu-1}$中必有不在 $E(T)$中的边.设 $f(T)=k$,即 $e_1,e_2,\cdots,e_{\nu-1}$在 T 与T^*上,而 e_k 不在T上,于是 $T+e_k$ 中有一个圈 C.令 e'_k 是在T上而不在T^*上的边,且 e'_k 在C上,显然,$T'=(T+e_k)-e'_k$ 也是生成树,又 $W(T')=W(T)+w(e_k)-w(e'_k)$.由算法知,$e_k$ 本是使$G[\{e_1,e_2,\cdots,e_k\}]$上无圈的权最小的边,又 $G[\{e_1,e_2,\cdots,e_k,e_{k'-1}\}]$是 T 之子图,也无圈,则有 $w(e'_k)\leqslant w(e_k)$.于是 $W(T')\geqslant W(T)$,即 T'也是最优树,但 $f(T')>k=f$

(T),与 $f(T)$之最大性矛盾,证毕.

例如求北京(Pe)、巴黎(Pa)、纽约(Ny)、伦敦(L)、墨西哥城(MC)与东京(T)六大城市间航线的最优连接.

用 Kruskal 算法求解的结果如图 2 所示,其中粗实线是求出的最优树上的边,此树总权为 122.

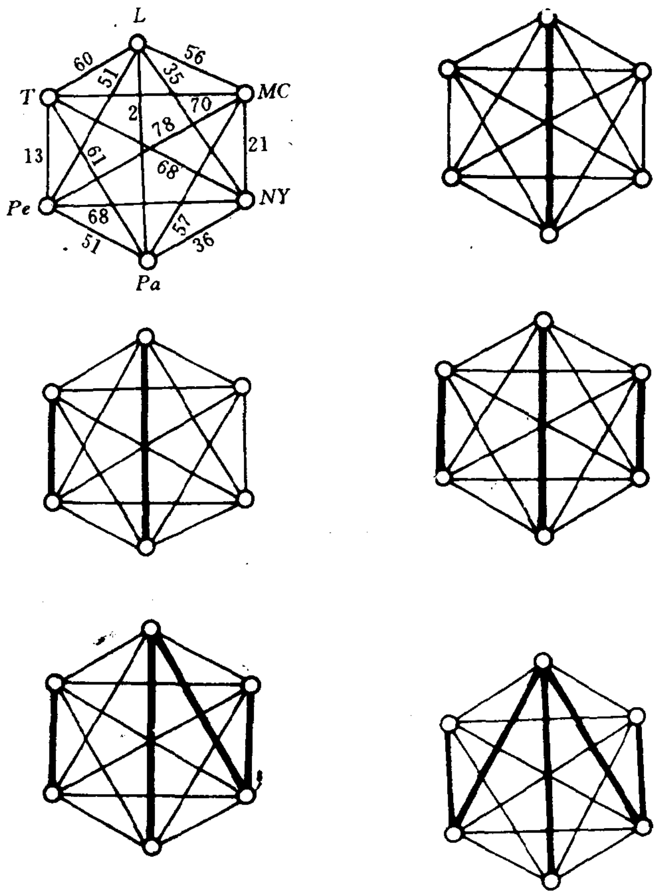


图 2

还有一个所谓 Steiner 连接问题:

在一个连通的铁路网上指定若干城市，要求从此铁路网上选出一个连通的子铁路网，使得上述指定的城市在此子网上，且此子网的总长度最小．

这个问题的数学模型是：

输入：连通图 $G(V,E)$，边权 $l(e)>0$，自然数 k，$X\subseteq V(G)$．

问：G 上是否有树 $T(W,F)$，使得 $X\subseteq W\subseteq V(G)$，$F\subseteq E(G)$，且 $\sum_{e\in F} l(e)\leqslant k$？

上述问题也称 Steiner 树问题，代号 ST，除了对特殊的图 $G(V,E)$，Steiner 树有有效算法外，至今尚未设计出对任何图 G 皆适用的有效算法！事实上，ST 是所谓 NPC 问题集合中的一元；NPC 是一个问题集合的代号，NPC 集合中的问题们至今无一个问题找到有效算法，而且已经证明，只要 NPC 中有一个问题有有效算法，则 NPC 中的每个问题都存在有效算法，反之，若能证明 NPC 中有一个问题确实不存在有效算法，则 NPC 中的每个问题都不存在有效算法．NPC 中的每个问题都是难以解决也许是真的不存在有效算法的问题！后面我们设专门的章节讨论这个问题．

1.3 碗摞问题

饭后，姊妹洗碗，妹妹把姐姐洗过的碗一个一个放在碗橱摞成一摞，共有 n 个图样两两相异的碗，洗前也摞成一摞，小妹贪玩，碗未必及时放入橱内，姐姐就把洗过的碗摞起来，问小妹摞起的碗可能有几种方式？

与上述摞碗问题有相同数学模型的问题还有许多，例如，一个汽车队在狭窄路面上行驶，不得超车，但可以进入一个死胡同去加油，之后再插队行驶，共 n 辆汽车，问可能有几种排列不同的车队开出城去？

在计算机存储问题当中，n 个字符都要进入“先入后出”存储器 S 恰一次，进入 S 时是有序的，问出 S 的不同字符列可能有几种？

上述三个问题的答案是一致的，为解决上述问题，我们引入“好括号列”的概念与其计数.

括号列是指由左括号“(”和右括号“)”组成的有限序列，所谓好括号列，是指

(1) 空列是好的.

(2) 若 A 与 B 是好括号列，则 AB 也是.

(3) 若 A 是好括号列，则 (A) 也是.

(4) 除(1)(2)(3)所称的好括号列外，再无其它好括号列(或称其它括号列为“坏括号列”).

例如(()(()))是好列，(()))(()是坏列.

定理 1 一个括号列是好括号列的充要条件是它由偶数个括号组成，其中一半是左括号，且从左向右读这个括号列时，读出的右括号个数不会超过读出的左括号个数.

证 若括号列是好列，显然它是由左括号占半数的偶数个括号组，下面用关于括号个数的归纳法证明从左到右读出的左括号个数不少于右括号读出的个数. 若括号数为 2，命题自真，设 m 个左括号和 m 个右括号的好列命题已真，我们考虑 n 个左括与 n 个右括组成的好括号列，其中 $m<n$.

(i) 若造此括号列时，最后一步是(2)，此括号列形如 AB，A 与 B 皆非空好列，从左向右读时，只要还在读 A，由归纳法假设，读出的左括不比右括少，当我们读到 A 的最后一个括号时，读出的左右括号个数一致，再读下去，即读 B，由归纳去假设，读出的右括号总数仍不超过读出的左括号总数.

(ii) 若造这个括号列时，最后一步是(3)，命题显然成立.

下面证明由左括号占半数的括号列，若从左至右读时，读出的左括号个数不少于右括号个数，则此括号列是好列. 仍用关于括号个数的归纳法，括号数为 2 时，命题显然成立，假设 $m<n$ 时，m 个左括号 m 个右括号组成的括号列命题已真，考虑 n 个左括 n 个右括号的括号列，从左向右读时，读了 $2m$ 个括号后，读得左右括号个数相等，

由归纳法假设,读出的这个子列是好括号列,右面未读的子列 B 也满足命题条件,由归纳法假设,B 也是好括号列,于是整个括号列 AB 也是好括号列.

若上述非空列 A 不存在,我们从左向右读时,读了第一个括号而未读其它括号时,由命题条件知第一个括号为左括,读到只剩一个括号未读时,已读出的左括号不比右括号少,又左括号右括号各占一半,故最后一个括号是右括号,于是原括号列为(A),A 仍满足命题条件,由归纳法假设,A 是好列,故(A)也是好括号列,证毕.

定理 2 由 $2n$ 个括号组成的好括号列总数为

$$C(n)=\frac{1}{n+1}C_{2n}^{n}\quad (C(n)\text{ 称为 Catalan 数}).$$

证 由 $p_1,p_2,\cdots,p_{2n}$是 n 个左括 n 个右括构成的坏括号列,由定理 1 知,有一个前缀,其中右括比左括多,设 $p_1,p_2,\cdots,p_j$ 的右括比左括多,且 j 最小,这时右括只比左括多 1 个,把从 p_{j+1}开始的每个括号"翻"过来,则得 $n-1$ 左括 $n+1$ 右括的坏括号列,显然这一变换是可逆的,故 n 个左括与 n 个右括组成的坏括号列与 $n-1$ 左括 $n+1$ 个右括组成的括号列一、一对应,而 $n-1$ 左括 $n+1$ 个右括组成的括号列共计 C_{2n}^{n+1}个,n 个左括 n 个右括组成的括号列共计 C_{2n}^{n}个,所以 $2n$ 个括号的好括号列共有:

$$\begin{aligned}C_{2n}^{n}-C_{2n}^{n+1}&=\frac{2n(2n-1)\cdots(2n-n+1)}{n!}-\frac{2n(2n-1)\cdots(2n-n)}{(n+1)!}\\&=\frac{1}{n+1}C_{2n}^{n},\end{aligned}$$

证毕.

回到我们前面提出的那三个实际问题,有字符进 S 时,记"(",出 S 时记")",当 n 个字符全部出了 S 后,得到了 n 个左括 n 个右括组成的 $2n$ 个括号的括号列,由定理 1,得到的是好括号列,又这种好括号列与出 S 的字符列之间一一对应,故我们的三个实际问题答案皆为 $C(n)$.

1.4 文章的码长问题

图 3 中画的是一棵以 v_0 为根的二元树(也称二叉树),兄弟关系左标 0 右标 1,每个叶表示一个字母,由根到叶的轨上的有序 0-1 序列为该叶的代码,00110110100001110111100010100001 的译文为 *Good morning*.

若设计一个 26 个叶的二元树,每个叶代表一个英语字母,则可以用此树的叶码表达任何一句话,进而表达任何一篇文章的信息.自然,我们希望总码长越短越好,字母(例如 Z)出现的频率小的,相应的叶码长一点还不怕,但字母出现的频率大的(例如 e)相应的叶码可不要太长,以期整个文章总码长最短.

这一问题的数学模型为所谓 Huffman 树.

以 v_0 为根 $v_1,v_2,\cdots,v_n$ 为叶的二元树中,轨 $P(v_0,v_i)(i=1,$

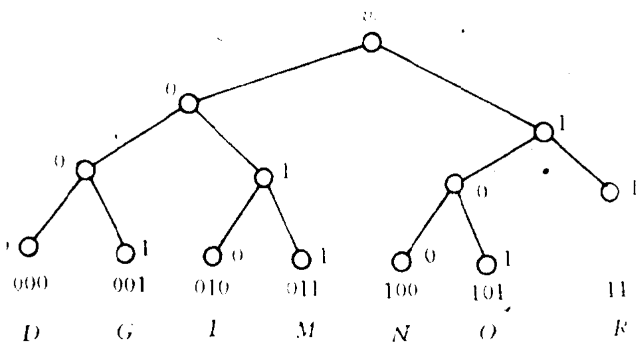


图 3

$2,\cdots,n)$的长 l_i 叫做 v_i 的码长,v_i 代表的事物出现的概率为 p_i,且 $\sum\limits_{i=1}^{n}p_i=1$,使得

$$m(T)=\sum_{i=1}^{n}p_il_i=\min$$

这种二元树为 Huffman 树,p_i 叫做权.

不妨设 $p_1 \leqslant p_2 \leqslant \cdots \leqslant p_n$.

定理 (1)T 是 Huffman 树,v_i, v_j 是兄弟,则 $l_i = l_j$.(2)v_1, v_2 是兄弟.(3)T^+ 是带权 $p_1 + p_2, p_3, \cdots p_n$ 的 Huffman 树,与 $p_1 + p_2$ 相应的叶子生出两个新叶分别带权 p_1 与 p_2,则所得树 T 是带权 $p_1, p_2, \cdots, p_n$ 的 Huffman 树.

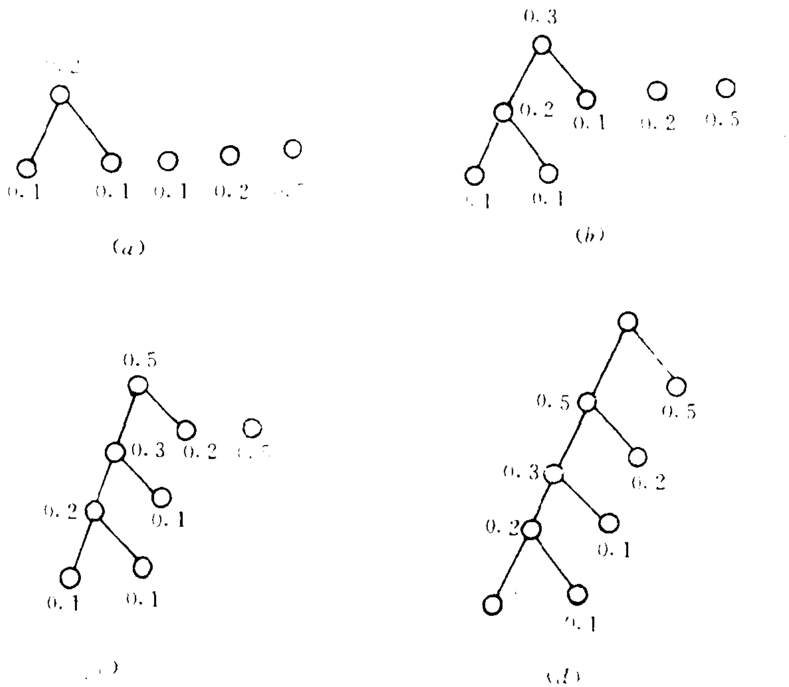


图 4

证 (1) 由树上轨 $P(v_0, v_i)$ 与 $P(v_0, v_j)$ 的唯一性,以及 v_0 到 v_i, v_j 之父的轨之唯一性,知 $P(v_0, v_i), P(v_0, v_j)$ 等长.

(2) 若 Huffman 树仅 v_1, v_2 两个叶,由 $m(T) = \min$ 知,v_1, v_2 是兄弟,码长 $l_1 = l_2 = 1$. 若 T 有三个以上的叶子,因 T 是二元树,故有码长不小于 2 的叶,又 $m(T) = p_1 l_1 + p_2 l_2 + \cdots + p_n l_n = \min$,不妨设 $l_1 \geqslant l_i, i = 2, 3, \cdots, n$. 这是因为,若 $p_k = p_1$,而 $l_k > l_1$ $(n \geqslant k \geqslant 2)$,我们把 v_1 与 v_k 足标对换,得一同构树,$m(T)$ 不变,而