

平面控制网网点优化排序研究*

王金岭

摘要 本文利用法方程系数阵的稀疏性,将 Banker 算法和 RCM 算法应用于平面控制网数据处理,实现网点自动优化排序,压缩了法方程系数阵的存贮空间。借助于各类平面网平差模型对 Banker 算法和 RCM 算法进行了比较、分析。采用优化排序算法不仅能扩大微机的解算容量,同时也提高了解算速度,减少了计算过程中的舍入误差,且适用于任意平面网、高程网、GPS 网的自动化数据处理。

关键词 平面网平差;稀疏矩阵;Banker 算法;RCM 算法;压缩存贮

分类号 P207 TP311.12

0 引言

采用电子计算机进行测量控制网数据处理时,法方程系数阵的存贮量最大,且与控制网中未知数个数的平方成正比。因此,在微机上解算大型测量控制网有一定困难。在微机上如何实现大型控制网的自动化数据处理是一个值得研究的技术问题。

事实上,由于平面控制网的边长、方向观测一般仅涉及少数位置相近的网点,使得法方程系数阵 N 具有稀疏性。因此,本文将稀疏矩阵技术应用于平面控制网的数据处理。在分析变带宽存贮法的基础上,采用稀疏矩阵技术对网点进行自动优化排序,显著地减少了 N 阵的存贮空间,扩大了微机的解算容量,提高了解算速度。

1 变带宽存贮技术

变带宽存贮技术(或称断面存贮技术)由英国学者 Jennings 首创于 1967 年,现已广泛应用于许多工程技术领域,在测量界也早已引起重视,并逐步引入到控制网数据处理软件中^[1,2]。下面借助于一个实例对变带宽存贮法进行分析。

如图 1 所示的测边网,有 2 个已知点(K_1, K_2), 25 个待定点,在平差计算中共计 50 个坐标未知数。在组成法方程之前,应首先确定未知数的顺序号。为处理方便起见,习惯上将每个网点的两个坐标未知数(dx, dy)排在一起,因此对未知数的排序实际上是对网点的排序。但是,计算过程中对网点编排的顺序号与网图上标明的实际点号(或点名)不同,一般由数据处理软件自动按待定点输入顺序编排^[3]。同一控制网,待定点的顺序号可能有多种,对图 1 所示的测边网按不同的方法给定了三种待定点的顺序,见表 1。

收稿日期:1992-10-08

* 国家教委优秀青年教师基金资助课题

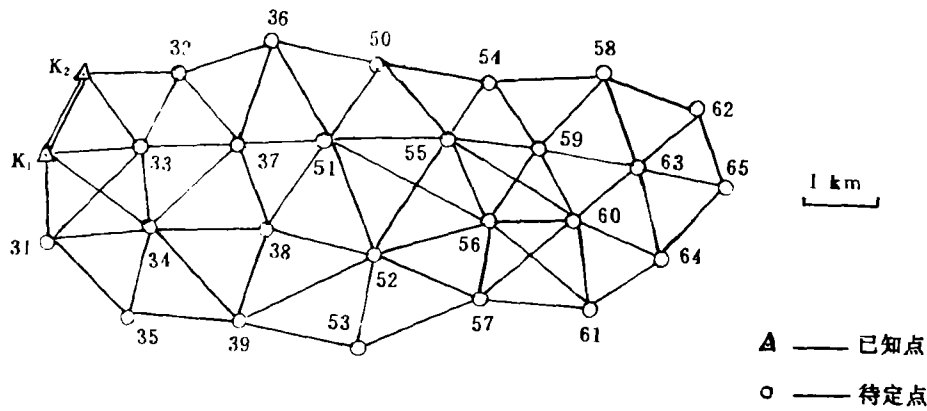


图1 测边网

待定点的顺序排定之后,法方程系数阵 N 的结构随之确定。图1所示的测边网与待定点随机的顺序列 A 对应的法方程系数阵 N 的结构如图2(a)所示。在 N 阵中,若第 k 列第一个非零元素为 N_{jk} ($j \leq k$),那么

表1 测边网网点排序结果

序号	序列 A	序列 B	序列 C
1	51	62	31
2	32	65	32
3	64	64	33
4	34	63	34
5	63	58	35
6	36	61	36
7	61	60	37
8	38	59	38
9	57	54	39
10	50	55	50
11	53	56	51
12	55	50	52
13	56	57	53
14	58	53	54
15	37	52	55
16	59	51	56
17	52	36	57
18	60	39	58
19	35	38	59
20	62	35	60
21	33	37	61
22	54	31	62
23	65	34	63
24	31	33	64
25	39	32	65
断面 P	1039	443	469

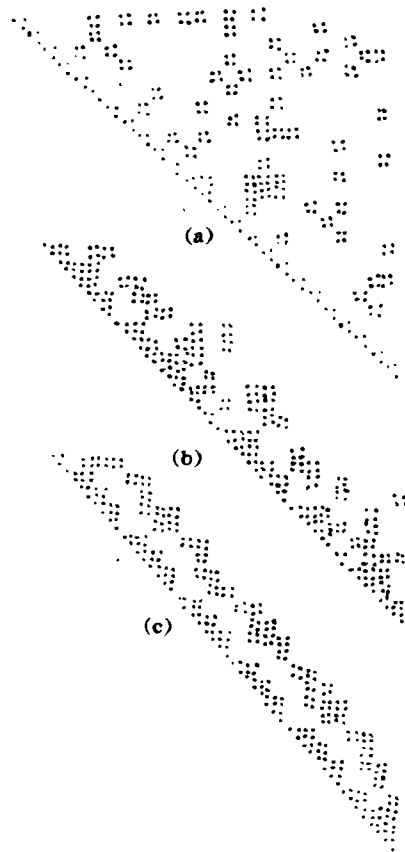


图2 法方程系数阵结构

$$S_i = k - j + 1$$

称为第 k 列的带宽, N 阵各列带宽之和称为断面, 若用 P 表示断面, 则 $P = \sum_{k=1}^n S_k$, n_k 为未知数个数。如图 2(a) 中第 3, 4 列的带宽分别为 1, 2; N 阵的断面 $P = 1039$ 。

显然, 带宽以外的元素为零, 并且在法方程约化的过程中永远保持不变, 不必参与运算, 也不需要存贮。所谓变带宽存贮法就是将法方程系数阵带宽内每一列对角线以上的所有元素一个接一个地按一维方式存贮在数组 AN 中, 每一列主对角线元素在 AN 中的地址存放在另一数组 ND 中, 那么 N 阵带宽内的元素与一维数组 AN 中的元素存在一一对应关系。由于带宽以外的零元素没有存贮, 这种方法一般来说总是优于上三角一维存贮。比如, 图 2(a) 中的 N 阵, 若用上三角一维存贮需要 1275 个元素, 而按变带宽存 1039 个元素, 减少了 236 个零元素, 但带宽内的零元素仍然多达 69%。

进一步分析, 我们注意到, 网点顺序号的改变(即未知数的重新排序)将导致 N 阵各列带宽 S_k 的变化, 合理的网点排序可以显著地提高变带宽存贮法的存贮效率。比如, 表 1 中的序列 B 是按优化排序算法(Banker)由计算机自动生成的, 与其对应的法方程系数阵 N 的结构如图 2(b) 所示, 断面仅 443, 其中零元素的比例由 69% 降为 21%。

由上述分析可知, 变带宽存贮法与上三角一维存贮法相比, 节省存贮空间是否显著还取决于网点的排序是否合理。传统的做法是按“相邻点号(与顺序号一致)差尽可能小”为原则, 人工对网点排序。表 1 中的序列 C 乃是人工精心排序的结果, 但其相应的法方程系数阵[如图 2(c) 所示]断面为 469, 还不如 Banker 算法的排序效果好, 并且对于较大规模的控制网, 人工优化排序难以实现。随着测绘技术的发展, 不仅测量控制网的规模在扩大, 而且数据采集逐步采用电子手簿, 数据处理朝着自动化方向发展^[3]。因此, 在平面测量控制网数据处理软件中, 有必要采用一套自动化的网点优化排序算法, 使得法方程系数阵 N 的断面极小化。

2 网点优化排序

在稀疏矩阵技术中, 有许多未知数优化排序算法可供选择, 其中由 Cuthill 和 Meke 于 1972 年提出的 RCM 算法^[2,5]以及 1976 年由 Snay 提出的 Banker 算法^[5]得到广泛应用。1986 年, 德国的 Benning 教授将 Banker 算法和 RCM 算法应用于地籍测量数据处理系统 KAFKA, 对未知数优化排序的效果进行了分析和研究^[4]。本文在简要介绍 RCM 算法和 Banker 算法的基础上, 重点讨论这两种算法应用于平面测量控制网数据处理的具体问题和实际效果, 并从中选择一个合适的排序算法应用于平面网数据处理。

2.1 RCM 算法和 Banker 算法简介

在讨论稀疏矩阵算法时, 一般是借助于“图”来表示稀疏矩阵的结构, 如图 3(a) 所示的对称阵可用图 3(b) 所示的“图”来描述。按照图论的定义, 一个节点的联接点个数称为该节点的“度”。

RCM 算法属于带宽极小化算法。其排序方法是, 从 n 个节点的图中具有最小度的某一节点 j 开始排序。假设已有 k ($1 \leq k \leq n$) 个节点被排序, 则从这 k 个节点中找出一个具有没被排序的联接点且为最小序号的节点 l ($l \leq k$), 将节点 l 的联接点中具有最小度的没排序节点排序号为 $k+1$ 。当 $k=n$ 时, 将上述顺序号倒排, 即为 RCM 算法的排序结果。

与 RCM 算法相比, Banker 算法较为复杂, 它是一种断面极小化算法。其排序步骤如下:

- (1) 从 n 个节点的图中选具最小度的某一节点 j , 排序号为 1, 并使 $V=j$;

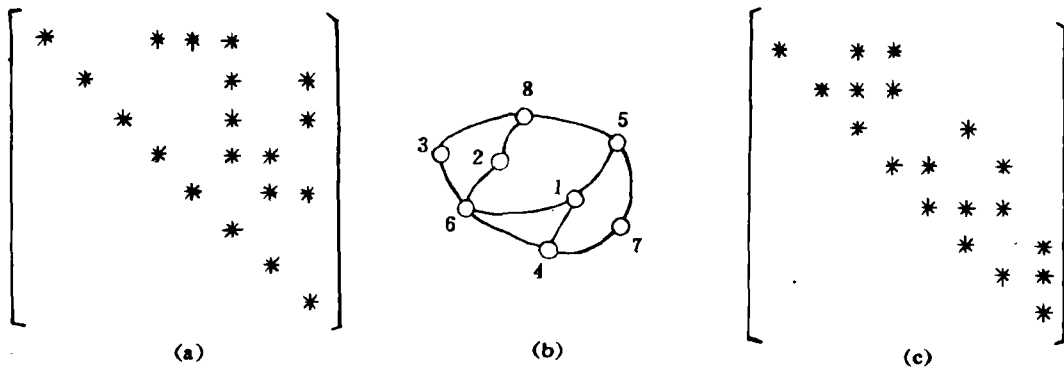


图3 矩阵结构示意图

- (2) V 的连接点中没排序号的节点存入(扩充)数组 A ;
- (3) 数组 A 中的所有节点以及它们的连接点中没排序号的节点存入(扩充)数组 B ;
- (4) 如果数组 B 不含任何节点, 排序结束;
- (5) 对数组 B 中的每一个节点计算其连接点中没排序且没存入数组 A 的节点个数 t ; 如该节点在数组 A 中 $r=1$, 否则 $r=0$; 若已有 $k(1 \leq k \leq n)$ 个节点被排序, 则数组 B 中与 $(t-r)$ 的最小值对应的节点 G 排序号为 $k+1$;
- (6) 从数组 A 或/和 B 中排除节点 G , 并使 $V=G$, 返回(2)。

按照上述 RCM 和 Banker 算法对图 3(b) 的节点重新排序, 其排序结果列于表 2。对于这个具体例子而言, 虽然两种算法的排序结果不同, 但压缩存贮空间的效果相同, 矩阵的断面由原来的 29 降为 22, 断面内的零元素由原来的 10 个降到 3 个, 与 Banker 算法排序结果对应的矩阵断面如图 3(c) 所示。事实上, 优化排序算法应用于较大规模的稀疏矩阵, 压缩存贮空间的效果更为显著。

表2 RCM 和 Banker 算法排序

节点	RCM	Banker
1	3	5
2	5	2
3	8	1
4	2	7
5	4	6
6	6	4
7	1	8
8	7	3

此外, 用上述算法排序时, 选择不同的起始点其排序法结果随之改变, 得到的断面亦不尽相同。Banker 算法的创立者 Snay 建议, 在图的最大直径的两端点附近各选 5 个节点分别做为起始点排序, 从中选择断面最小的排序结果。在本文的研究中采用了这种改进措施。

2.2 网点优化排序方案设计

RCM 算法和 Banker 算法优化排序的对象都是未知数, 因而需要建立所有未知数之间的联接矩阵, 以便将法方程系数阵中非零元素的结构用相应的“图”来描述。由于坐标未知数之间的联系一般由观测值的类型确定, 在组成法方程之前可以根据输入的观测值建立未知数的联接矩阵。然而, 在平面控制网中, 消去定向角未知数之后, 每个点仅有两个坐标未知数, 无论边长或方向观测值均同时与网点的坐标未知数有关。所以, 一个网点的两个坐标未知数的“度”相同(习惯上也总排在一起), 由网点代替其坐标未知数排序, 其效果是相同的。而由网点建立的联接矩阵与坐标未知数建立的联接矩阵相比, 节省一半的存贮空间, 同时还可以加快排序的速度。对于三边网而言, 由网点建立的联接矩阵与文献[3]提出的网点数据结构一致; 当控制网中有方向观测值时, 由于消去定向角未知数引入和方程, 这种联接矩阵中还应包括相应的网点信息。

此外,为编程方便起见,网点排序时不区分已知点或待定点,排序结束之后,再按待定点序号大小形成一个连续的待定点顺序序列。

2.3 网点优化排序的效果与分析

根据上述分析,作用分别编制了用 RCM 算法和 Banker 算法进行网点优化排序的软件,并将其扩充到 APS-H 软件系统^[3]中,应用于各类实测网,模拟网的数据处理,其中与网点优化排序效果有关的数据列于表 3 中。

表 3 网点优化排序结果

平面控制网(名)	N_1	\bar{r}	N_2	上三角一维存贮		变带宽存贮(Jennings)			RCM+Jennings			Banker+Jennings		
				P_{Δ}	$\frac{N_2}{P_{\Delta}}(\%)$	P_J	$\frac{P_J}{P_{\Delta}}(\%)$	$\frac{N_2}{P_J}(\%)$	P_R	$\frac{P_R}{P_J}(\%)$	$\frac{N_2}{P_R}(\%)$	P_B	$\frac{P_B}{P_J}(\%)$	$\frac{N_2}{P_B}(\%)$
A. 三边网	50	0.24	319	1275	25	1039	81	31	443	43	72	443	43	72
B. 混合网	118	0.36	1057	7021	15	4257	61	25	2233	52	47	2189	51	48
C. 三角网	148	0.47	2446	11026	22	9806	89	25	4986	51	49	4498	46	54
D. 图根导线	214	0.14	1377	23005	6	6581	30	21	4877	74	28	2657	40	52
E. 城市导线	582	0.14	3585	169653	2	37157	22	10	22277	60	16	10889	29	33
F. NET(25, 20, 0)	992	0.48	12516	492528	3	170448	35	7	61428	36	20	63708	37	20
H. NET(25, 32, 5)	1592	0.14	11528	1268028	1	109844	9	10	111624	102	10	43648	40	26
I. NET(25, 40, 5)	1992	0.14	14464	1985028	1	140012	7	10	147984	106	10	54236	39	27
J. NET(29, 36, 6)	2080	0.12	15836	2164240	1	145520	7	11	168788	116	9	55612	38	28

注: N_1 : 未知数个数, 不含消去的定向角未知数; \bar{r} : 平均多余观测分量, 即网中多余观测数与观测值个数之比;

P_{Δ} : 法方程式系数阵 N 的存贮元素个数; N_2 : 法方程式系数阵 N 带宽内的非零元素个数。

分析表 3 中的数据之前,先介绍一下表 3 中作为算例的几个控制网。网 A 为图 1 所示的测边网,网 B 为图 4 所示的混合网(作者为研究控制网数据处理自动化而建立的试验网)。网 C 是某城市三等三角网,9 个已知点,74 个待定点。网 D 为海南某测区数字化测图生产中布设的基础图根网,其中有 6 个已知点,107 个待定点,由 DH-PR1 电子手簿采集数据,采用 APS-H 软件系统^[3]实现数据处理自动化。网 E 为某城市一个区的一级导线网,22 个已知点,291 个待定点。表 3 中模拟网 F, H, I, J 的构成方式为:

假设地面上有 $m \times n$ 个网点等间距构成矩形点阵,点阵的四周以及每一排横向的点之间布设了导线,四个顶点为已知点。例如,当 $m=5, n=4$ 时,网形如图 5(a) 所示。为了加强网的结构,可以在点阵的纵向方向等间距地加测方向和边长,从而构成结点稀密不同的各种控制网。设两条纵向导线结点之间的过渡点数为 ID ,那么,按上述方式构成的网可用一个函数 $NET(m, n, ID)$ 来描述。例如,与 $NET(5, 4, 1)$ 对应的控制网如图 5(b) 所示。与 $NET(5, 4, 3)$ 对应的控制网如图 5(a) 所示。显然,当 $ID=0$ 时,所有待定点均为结点,所以, $NET(25, 20, 0)$ 实际上是 500 个网点方格网,而 $NET(25, 32, 5)$, $NET(25, 40, 5)$ 和 $NET(29, 26, 6)$ 则分别为 800, 1000 和 1044 个网点的导线网。

分析表 3 中的数据,可以得出以下几点结论:

① 随着网点数的增加,法方程系数阵中的非零元素所占的比例显著减小。网点数超过 100 的情况下,非零元素不足 10%,法方程系数阵是典型的稀疏矩阵。

② 与上三角一维存贮法相比,变带宽存贮法的存贮效率高。最有利的情况下,如模拟网 $NET(25, 40, 5)$,变带宽存贮量仅占上三角阵存贮量的 7%。从表 3 中的数据可以看出,如果不对网点进行优化排序,仅采用 Jennings 变带宽存贮法在微机上难以解算超过 496 个待定点的模拟网 F, H, I, J, 而采用 Banker 算法对网点进行自动优化排序之后则能够在微机上解算。

③ 无论是测边网,三角网或是方格网,导线网,网点优化排序显著地提高了变带宽存贮法的存贮效率。以 Banker 算法为例,网点优化排序之后的法方程式系数阵 N 之断面 P_B 与网点

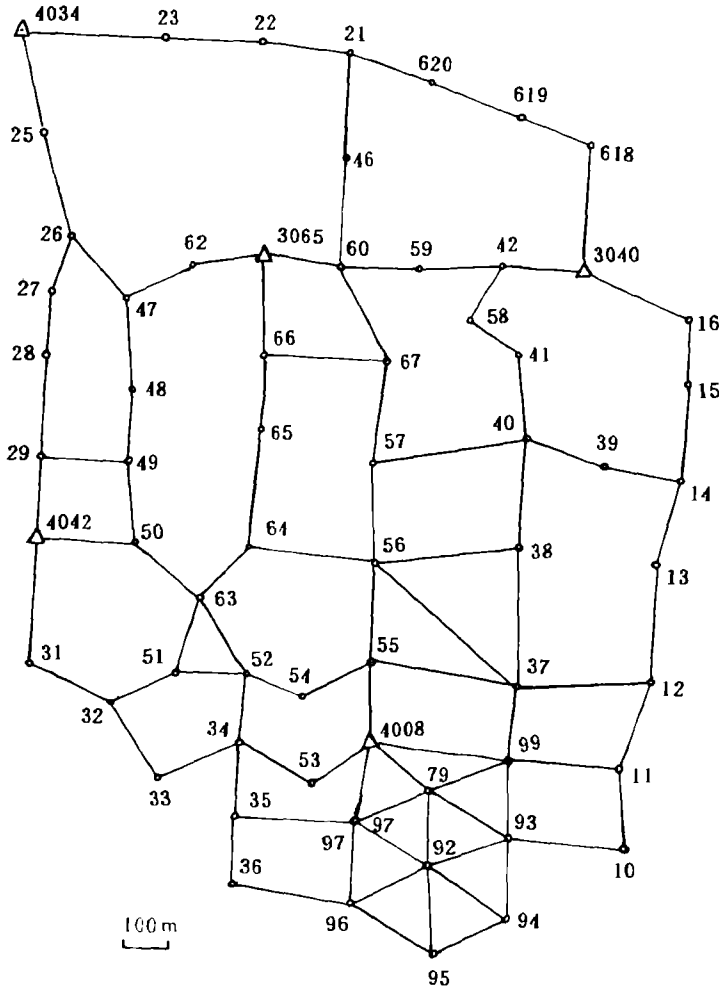


图 4 混合图网

未优化排序断面 P_i 相比,压缩存储空间大都在 50% 以上。表 3 中三角网 (C), 压缩空间达 54%; 城市导线网 (E), 压维空间高达 71%; 几个模拟网 (F, H, I, J), 压缩存储空间也多达 60% 以上, 且方面没有超出软件 APS-H 中给定的存储空间。

④ 多数情况下, Banker 算法优于 RCM 算法, 尤其是控制网比较稀

疏, 平均多余观测分量 \bar{r} 较小的情况下, 如表 3 中控制网 D, E, H, I, J, Banker 算法的优势更显著。当 $\bar{r} \geq 0.35$ 的情况下, 如控制网 A, B, C, F 等, Banker 算法与 RCM 算法的排序效果接近。

此外, 采用稀疏矩阵技术还有利于加快解算的速度。例如, 在 GW 微机 (80286CPU) 上采用

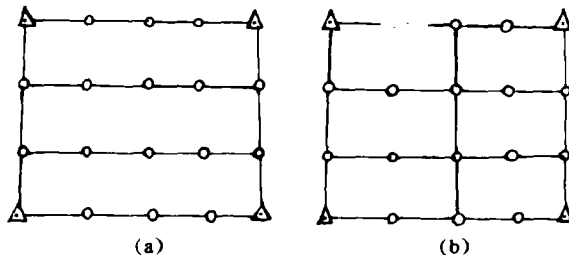


图 5 等距矩形点阵网

APS-H 软件处理城市导线网 E , 不对网点优化排序的情况下法方程解算与 N 阵求逆共需要 65 分钟; 若采用 Banker 算法对网点进行优化排序则仅需要 30 分钟, 而网点优化排序的时间还不足 2 分钟, 在解算时间方面提高效率达 50%。

采用稀疏矩阵技术, 还大大减少了法方程解算过程中不必要的运算操作, 减少了舍入误差, 保证大型平面控制网解的稳定性, 提高解算的精度。例如, 在 TMS 微机 (80486CPU) 上解算多达 1992 个坐标未知数的模拟网 $NET(25, 40, 5)$, 解算时间仅 105 分钟, 平差后的观测值、坐标值与相应的理论值一致, 误差小于 0.01mm (或秒), 所有观测值的多余观测分量之和与整个网的多余观测数相等, 计算误差不到 0.001。由此可见, 计算精度很高。

致谢 在本文研究过程中, 作者先后得到张正禄教授、吴俊昶教授的鼓励和指教, 特此致谢!

参 考 文 献

- 1 吴俊昶, 刘大杰. 控制网测量平差. 北京: 测绘出版社, 1985. 358~365
- 2 杨绍祺等. 稀疏矩阵——算法及其程序设计. 北京: 高等教育出版社, 1985. 359~386
- 3 王金岭. 平面测量控制网数据处理自动化软件系统的研究和实现: [学位论文]. 武汉测绘科技大学工测系, 1989. 52~58
- 4 Benning W. Analyse Hybrider Lageaufnahmen in Sparse-Technik. zfv, 1986, 111(11); 506~513
- 5 Snay R. A. ; Reducing the Profile of Sparse Symmetric Matrices. Bulletin Geodesique. 1976, 50(4); 341~352

The Research on the Optimal Sequence of Net Points in Horizontal Network Adjustments

Wang Jinling

Abstract In this paper, the necessity and realization of sparse matrix-techniques in adjustments of horizontal networks are discussed. Some numerical tests indicate that in many cases Banker algorithm can produce lower profiles than RCM algorithm. With sparse-matrix techniques, adjustment programs can not only calculate larger networks on micro-computers but also speed up calculating and reduce the round-off errors in adjustment results.

Key words adjustments of horizontal networks; sparse-matrix; Banker algorithm; RCM algorithm; compress storage