最优化方法的理论及应用

（三）

—最优化方法介绍—

中国科学院数学研究所运筹室编

1977.1

说　明

本资料主要介绍最优化方法，并力图用通俗的方式描述方法的思路，为希望了解一些最优化方法的实际工作者提供一个参考。因此，方法的收敛性及收敛速度等问题的证明均不涉及。

关于一维最优化的方法，这里不再叙述，可参考其他有关资料。

目　　录

第一章 直接法

§1. 因素轮换法

因素轮换法又称坐标转换法是用得比较多的一个方法。这个方法不仅数学工作者在计算工作中常常采用，许多实际工作者在试验工作中也经常采用，简便易行。计算步骤如下。

§1.1 两个因素的情形

设两个因素分别用 x_1，x_2 代表，$x=(x_1, x_2)$，$f(x)=f(x_1, x_2)$ 为目标函数，要求 $f(x_1, x_2)$ 在 $a_1 \leq x_1 \leq b_1$，$a_2 \leq x_2 \leq b_2$ 上的极大值。当因素 x_1，x_2 分别固定于给定值 $x_1^{(0)}$，$x_2^{(0)}$ 时，就给出了一个试验方案 $x^{(0)}=(x_1^{(0)}, x_2^{(0)})$ 并称在 $0x_1x_2$ 平面上给出了一个"点"。

因素轮换法就是从一个初始的试验点出发按因素 x_1，x_2 逐个地进行一个因素的优选。以两次优选后得到的点作为新的初始点，重复上述过程，直到达到精度要求时为止。作法如下：

若 $x^{(0)}=(x_1^{(0)}, x_2^{(0)})$ 为给定的初始点，从 $x^{(0)}$ 出发，固定因素 $x_2=x_2^{(0)}$，对因素 x_1 进行优选，即

求
$$\max_{a_1 \leq x_1 \leq b_1} f(x_1, x_2^{(0)})$$

若最大值在 $x_1=x_1^{(1)}$ 时达到，即

$$\max_{a_1 \leq x_1 \leq b_1} f(x_1, x_2^{(0)})=f(x_1^{(1)}, x_2^{(0)})$$

记 $A_1=(x_1^{(1)}, x_2^{(0)})$，也就是说，在过 $x^{(0)}$ 的水

平直线上求$f(x)$的极大值点得到A_1。显然，A_1不会比$x^{(0)}$差。从A_1出发（也就是固定$x_1=x_1^{(1)}$对第二个因素x_2进行优选，即在过A_1点的竖直线上求$f(x_1,x_2)$的极大值，若

$$\max_{a_2\le x_2\le b_2} f(x_1^{(1)},x_2)$$

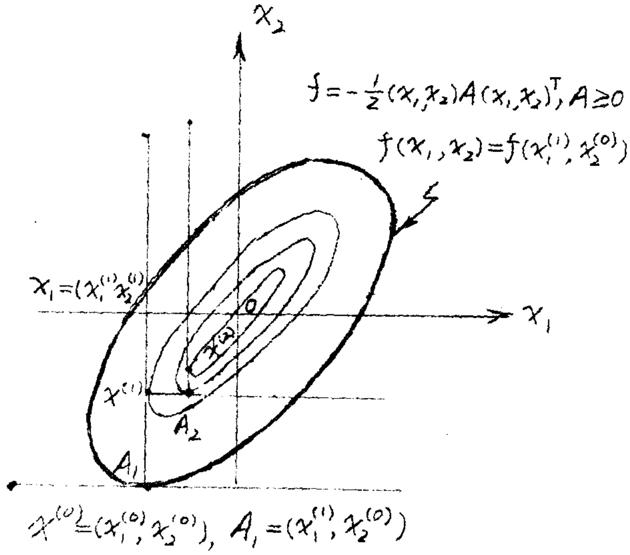


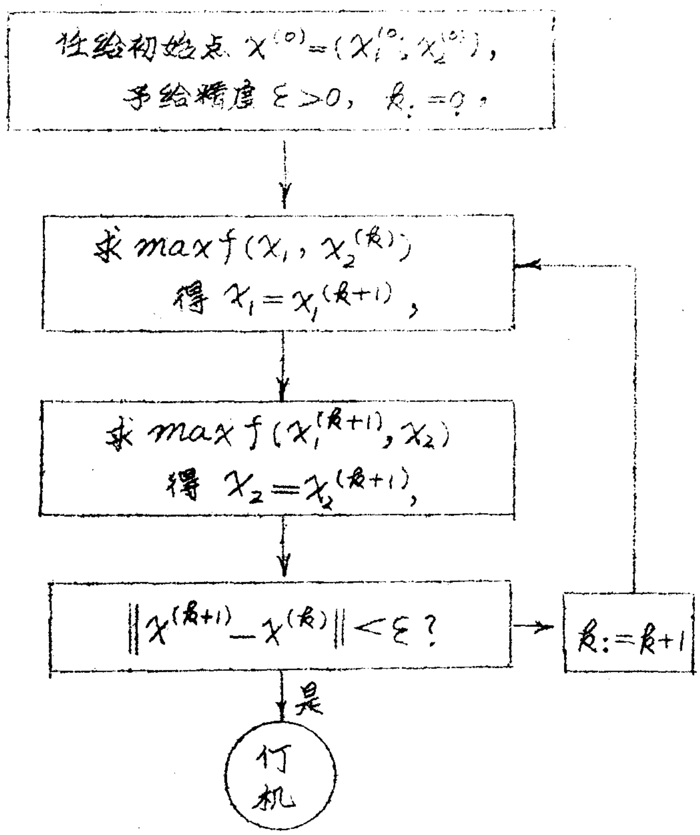
图 1.1

在$x_2=x_2^{(1)}$时达到，即

$$\max_{a_1\le x_2\le b_2} f(x_1^{(1)},x_2)=f(x_1^{(1)},x_2^{(1)}),$$

此时令$x^{(1)}=(x_1^{(1)},x_2^{(1)})$，同样，$x^{(1)}$不会比$A_1$差，到现在为止，两个因素$x_1$，$x_2$都已轮流优选过一次，也就是在水平与竖直两个方向上都作过一次一维（即一个因素）的最优化，如图）。若予给精度$\varepsilon>0$且$x^{(1)}$与$x^{(0)}$的距离$\|x^{(1)}-x^{(0)}\|<\varepsilon$，则取$x^{(1)}$为最优解；否则以$x^{(1)}$为新的初始点重复以$x^{(0)}$为起点先沿水平后沿竖直两个方向的逐次优选的过程，依此循环往复，直至达到满足精度要求时为止。

计标框图



§1.2 n个因素的情形

n维的情形与二维情形完全类似。由于此时有n个因素，故一个方案对应于一个n维向量 $x=(x_1, x_2, ----x_n)$ 且一个因素的优选在每一轮中就需要作n次。具体作法如下。

若初始点为 $x^{(0)}=(x_1^{(0)}, x_2^{(0)}, ----x_n^{(0)})$，从 $x^{(0)}$ 出发，将因素除 x_1 外的其他 $n-1$ 个因素 $x_2, x_3 ----x_n$ 暂时固定于 $x_2^{(0)}, x_3^{(0)}----x_n^{(0)}$，对第一个因素 x_1 进行优选，即求

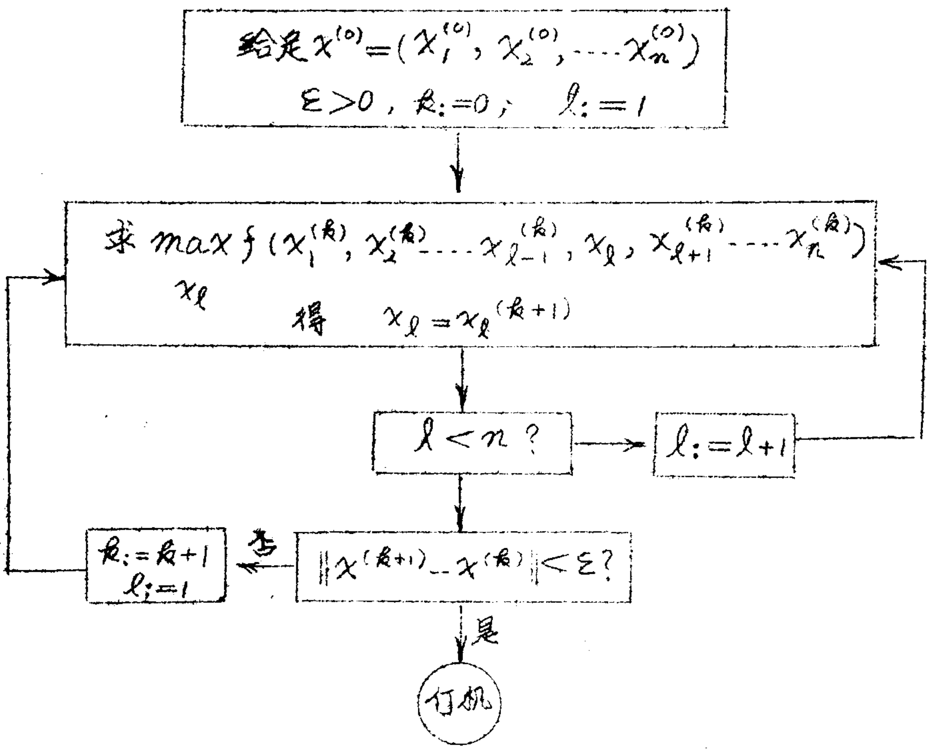
$$\max_{x_1} f(x_1, x_2^{(0)}, x_3^{(0)} -----x_n^{(0)}) \quad ,$$

若$f(x)$当$x_1=x_1^{(1)}$时达到最大值，记$A_1=(x_1^{(1)},x_2^{(0)}----x_n^{(0)})$，然后以$A_1$为起点，令$x_1=x_1^{(1)}$，$x_3=x_3^{(0)}$，$x_4=x_4^{(0)}----x_n=x_n^{(0)}$，对因素$x_2$进行优选。即求

$$\max_{x_2} f(x_1^{(1)},x_2,x_3^{(0)},x_4^{(0)}-----x_n^{(0)})$$

若极大值在$x_2=x_2^{(1)}$时达到，记$A_2=(x_1^{(1)},x_2^{(1)},x_3^{(0)},x_4^{(0)},---x_n^{(0)})$并以$A_2$为起点令$x_1=x_1^{(1)}$，$x_2=x_2^{(1)}$，$x_4=x_4^{(0)}$，$----x_n=x_n^{(0)}$对因素$x_3$进行优选，若优选后得点$A_3=(x_1^{(1)},x_2^{(1)},x_3^{(1)}x_4^{(0)},----x_n^{(0)})$，以$A_3$为起点依此方式对第四个因素进行优选，这样按因素的顺序进行的优选进行$n-1$次后得点$A_{n-1}=(x_1^{(1)},x_2^{(1)},----x_{n-1}^{(1)},x_n^{(0)})$，再以$A_{n-1}$为起点，固定$x_1=x_1^{(1)}$，$x_2=x_2^{(1)}$，$----x_{n-1}=x_{n-1}^{(1)}$对因素$x_n$进行优选，最后得到点$A_n=(x_1^{(1)},x_2^{(1)},----x_n^{(1)})$，到现在为止，一个因素的优选依此作了$n$次，令$A_n=x^{(1)}$，显然，上述$A_{i+1}$点不会比$A_i$点差。以$x^{(1)}$为新的初始点，重复上述逐次的优选过程。若某一优选的起点为$x^{(k)}$，以$x^{(k)}$为起点。分别对$x_1,x_2,----x_n$的n次优选后的终点为$x^{(k+1)}$，$x^{(k)}$与$x^{(k+1)}$满足$\|x^{(k+1)}-x^{(k)}\|<\varepsilon$，此时取$x^{(k+1)}$为最优解；否则又以$x^{(k+1)}$为新的初始点再重复上述过程，依此循环往复，直至无法进行或达到精度要求为止。

计祘框图



§1.3 定步长轮换法

上面的因素轮换法是顺序沿每一个坐标轴方向作函数的一维最优化，这种手续也可以用定步长的轮换法来代替，在实际计祘中，这种作法也能得到很好的效果。步骤如下（参看步长加速法）：

先选定步长 $\delta>0$ 和初点 A，予给精度 $\varepsilon>0$，若坐标轴方为 $e_1=(1,0,\cdots 0)$，$e_2=(0,1,0,\cdots 0)$，$\cdots e_n=(0,\cdots 0,1)$，计祘程序如下：

1). $Z:=A$，$x_0:=A$．

2). $k:=1$，计祘 $f(x_{k-1}+\delta e_k)$．

若 $f(x_{k-1}+\delta e_k)>f(x_{k-1})$，令 $x_k=x_{k-1}+\delta e_k$，转至 3)；

若 $f(x_{k-1}+\delta e_k)\le f(x_{k-1})$，计祘 $f(x_{k-1}-\delta e_k)$，

若 $f(x_{k-1}-\delta e_k)>f(x_{k-1})$，令 $x_k=x_{k-1}-\delta e_k$ 转至3)；

若 $f(x_{k-1}-\delta e_k)\leq f(x_{k-1})$，令 $x_k=x_{k-1}$，　转至3)；

3) $k<n$？是，$k:=k+1$ 转至2)；

否，转至4)；

4). $f(x_n)>f(z)$？是，则 $z:=x_n$，$x_0:=x_n$，$k:=1$ 转至2)；

若否，则问 $\delta<\varepsilon$？

若 $\delta<\varepsilon$，停机．

否则，$\delta:=\frac{1}{2}\delta$，

$z:=x_n$，$x_0:=x_n$，$k:=1$，

转至2)；

注：记号 $X:=Y$，若 X，Y是数，表示 X取Y值；若 X，Y 是向量，表示 X 的各分量取 Y 的对应分量的值。

§2. 步长加速法

我们知道，瞎子爬山法、座标轮换法等方法是从某一点开始，在其周围进行探测，若沿某一方向探测成功，则向前移动一步，再以这点为新的起点开始新的探测，若不成功，则不移动，改变探测方向，再行探测。这种手续重复进行，以求一步一步的登高。但是，这样的作法不考虑行进的路线，也就是不考虑从前面某一个成功的点起走到现在这个高程是沿着什么方向走过来的。如果注意到这一点，就使我们有可能利用前面的信息以加速行进的步伐。人们登山的时候，沿着山脊往上走可以爬到峰尖。步长加速法就是以此为背景而进行设计的。目的是边走边看，探测有利的方向，使行进的路线与山脊不会有太大的偏离，加速登山的进程。

这个方法略为复杂，但实践证明，效果较好，且适于在电子计标机上使用。

计标方法

这个方法由两个步骤组成，两个步骤交替地进行。这两个步骤是：

I. 在某一点周围进行局部探测；

II、沿探测得到的上升方向的平移；

前者是为了求得一个上升方向，后者是为了加大步伐前进。

为明了起见，先讲两个变数的情形，后讲 n 个变数的情形，后者只是符号复杂一点而已。

§2.1 二维的标法

以 $x=(x_1, x_2)$ 为两个变数 x_1, x_2 组成的二维向量，称为平面上的点。今要求目标函数 $f(x)=f(x_1, x_2)$ 的极大值，以 $e_1=(1,0)$, $e_2=(0,1)$ 表两个互相垂直的坐标轴的方向，$\delta_1>0$ 为沿 e_1 方向探测时的步长，$\delta_2>0$ 是沿 e_2 方向探测时的步长，数值大小酌情选定，计标步骤如下：

1、在某一点周围的局部探测：

先给定初始点 (x_1^0, x_2^0)。在 (x_1^0, x_2^0) 的周围依次按 e_1, e_2 方向进行探测。在探测过程中，除沿座标方向探测得到的中间点外，局部探测的终点我们称为"基点"，以 B 记之，平移后得到的点我们称为"参考点"，以 R 记之，并规定初始点是基点又是参考点，记 $(x_1^0, x_2^0)=R_0=B_0$.

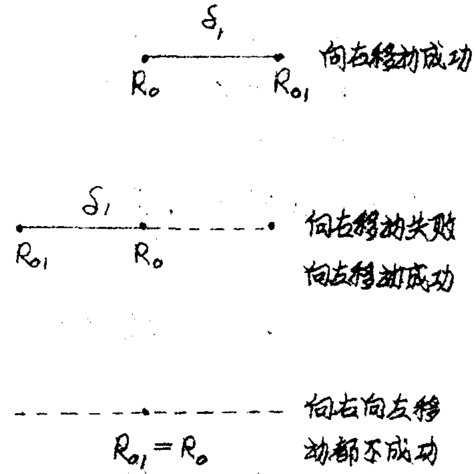
从 R_0 出发，先沿 e_1 方向以 δ_1 为距离向右移动一步，得点 $R_0+\delta_1 e_1=(x_1^0, x_2^0)+\delta_1(1,0)=(x_1^0+\delta_1, x_2^0)$。若比较函数值时

出现 $f(R_0+\delta_1 e_1)>f(R_0)$，称移动"成功"，记得到的点为 R_{01}，此时，$R_{01}=R_0+\delta_1 e_1$；若 $f(R_0+\delta_1 e_1)\leq f(R_0)$，说明向 e_1 的正方向移动"失败"，退回 R_0。从 R_0 出发，再向左即向 e_1 的反方向以 δ_1 为距离移动一步，进行试探，得点 $R_0-\delta_1 e_1$，若 $f(R_0-\delta_1 e_1)>f(R_0)$，则移动"成功"，令 $R_{01}=R_0-\delta_1 e_1$，否则令 $R_{01}=R_0$。

水平探测的结果有三种可能存在（如图 1.2）：'向右移动成功'，'向右移动失败向左移动成功'、向右向左移动都不成功。三者必居其一。按照三种可能结果，R_{01} 的坐标可如下表示；若记 $R_{01}=(\tilde{x}_1,\tilde{x}_2)$，则

$$\tilde{x}_1=\begin{cases} x_1^0+\delta_1 & \text{若向右移动成功；} \\ x_1^0-\delta_1 & \text{若向右移动失败向左移动成功；} \\ x_1^0 & \text{若向左向右均失败，} \end{cases}$$

$$\tilde{x}_2=x_2^0\ 。$$

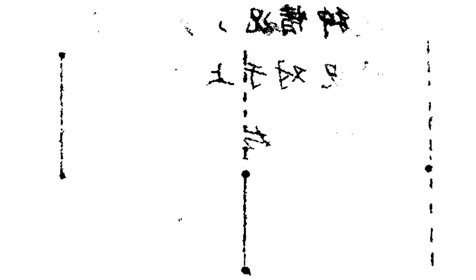


（图 1.2）

由于移动是在水平方向进行，第二个坐标不变。不管是出现那一种情形，最后总可确定一个点 R_{01}，沿 e_1 的探测也暂告结束。

水平方向即沿 e_1 方向的探测结束以后，从 R_{01} 出发开始沿方向 e_2 进行探测。从点 R_{01} 出发，以 δ_2 为距离先向上试探一步，得点 $R_{01}+\delta_2 e_2=(\tilde{x}_1,\tilde{x}_2)+\delta_2(0,1)=(\tilde{x}_1,\tilde{x}_2+\delta_2)$，若 $f(R_{01}+\delta_2 e_2)>f(R_{01})$，则移动'成功'，记得到的点为 R_{02}，

$R_{02}=R_{01}+\delta_2 e_2$。反之，称为‘失败’，退回 R_{01}，然后再向下以 δ_2 为距离移动一步，得点 $R_{01}-\delta_2 e_2=(\tilde{x}_1,\tilde{x}_2-\tilde{\delta}_2)=$ $=(\tilde{x}_1,x_2^0-\delta_2)$。若 $f(R_{01}-\delta_2 e_2)$ $\geq f(R_{01})$，则探测‘成功’，记此点为 R_{02}，$R_{02}=R_{01}-\delta_2 e_2$。否则即为‘失败’，退回 R_{01}。



（图1.3）

与 e_1 方向的探测类似，从 R_{01} 出发的沿方向 e_2 的探测将会出现‘向上移动成功’、‘向上移动失败向下移动成功’、‘向上向下移动均失败’这三种可能的结果，三者必居其一，（如图1.3）。按照这三种可能情形，若记 $R_{02}=(\hat{x}_1,\hat{x}_2)$，则 R_{02} 的坐标可表示为：

$$\hat{x}_1=\tilde{x}_1$$

$$\hat{x}_2=\begin{cases} x_2^0+\delta_2 & \text{若向上移动成功;} \\ x_2^0-\delta_2 & \text{若向上移动失败向下移动成功;} \\ x_2^0 & \text{若向上向下移动均失败;} \end{cases}$$

显然，经过先沿 e_1 后沿 e_2 的探测以后，最后总可得到一个点 R_{02}。这个点 R_{02} 就是从 R_0 出发在 R_0 周围进行局部探测后得到的点，至此在 R_0 周围的局部探测即告完成。局部探测得到的点我们称为‘基点’，记 R_{02} 为 B_1。

局部探测的结果得到一个新的基点。对于这个新的点可分二种情况：新的基点不比前面一个基点优越，即 $f(B_{i+1})\leq f(B_i)$，B_i 表示过程中第 i 次确立的基点，B_{i+1} 为接续产生的基点，

碰到这种情况，新的基点 B_{i+1} 不能确立，弃之并退回至 B_i。这种情况对于上面的探测来说就是 $B_1=B_0$；$(B_0=(x_1^0, x_2^0))$ 也就是沿左、右、上、下四个方向探测均"失败"，没有找到比 B_0 更好的点，此时 B_1 不能确立，基点仍是 B_0。如果从 B_0 出发探测以后就碰到这种情况，由于已经用步长 δ_1，δ_2 探测过一次，此时要分别以 $\frac{1}{2}\delta_1$，$\frac{1}{2}\delta_2$ 代替 δ_1 与 δ_2，重新施行步骤I。若重复多次以后仍出现这种情况，且步长已缩小到予给的精度，则视此点为最优解。一般说来，这种情况只有在过程进行相当阶段以后才会出现。第二种情况是局部探测成功，即 $f(B_{i+1})>f(B_i)$，此时，$B_{i+1}\neq B_i$ 即 B_{i+1} 与 B_i 的坐标不全相等，方向 $B_{i+1}-B_i$ 就视为从 B_i 出发的上升方向。对于上面的情况，上升方向就是 B_1-B_0。因此时有 $B_1\neq B_0$，$f(B_1)>f(B_0)$。

2、沿上升方向的平移：

局部探测成功以后，确立了一个新的基点，同时也确立了一个新的上升方向。上升方向找到以后接着进行步骤II。平移就是从新确立的基点出发在新、旧二个接续的基点的联线上进行，平移的距离等于这两个接续的基点的距离，这样又得到一个新的点。对于上面的情形，从 B_1 出发在 B_1-B_0 的方向上平移后得到的点为：

$$B_1+(B_1-B_0)=2B_1-B_0=2(\hat{x}_1,\hat{x}_2)-(x_1^0, x_2^0)=(2\hat{x}_1-x_1^0, 2\hat{x}_2-x_2^0)$$

B_1-B_0 表示从 B_0 指向 B_1 的向量。平移后得到的这个点记为 R_1，$R_1=B_1+(B_1-B_0)=2B_1-B_0$。一般情形，我们就得到点 R_{i+1}，$R_{i+1}=B_{i+1}+(B_{i+1}-B_i)=2B_{i+1}-B_i$。至此，平移的步骤即告完成。

从一点出发沿某一方向平移后得到的点我们称为"参考点"。

显然，这种大踏步的移动不一定成功，成功与否不视 $f(R_{i+1})>f(B_{i+1})$ 与否而定，而是再从 R_{i+1} 出发施行步骤I得到一个新基点 B_{i+2} 后，视 $f(B_{i+2})>f(B_{i+1})$ 与否而定。若 $f(B_{i+2})>f(B_{i+1})$，即平移以后再进行局部探测得到一个优于原基点的新基点，此时确定 B_{i+2} 并称平移"成功"。对于上面的情形，平移后得到 R_1，从 R_1 出发施行步骤I得 B_2，若 $f(B_2)>f(B_1)$，则平移成功，反之则失败退回 B_1，这时从 R_1 出发的探测工作因无效而放弃了。由于平移成功后确立了一个新的基点，则由这个新基点出发又可再进行平移，得到一个新的参考点；如果失败，则退回至原来的基点而重新进行步骤I，这时，这个原来的基点就视为新的初始点，也是目前得到的最好的点。

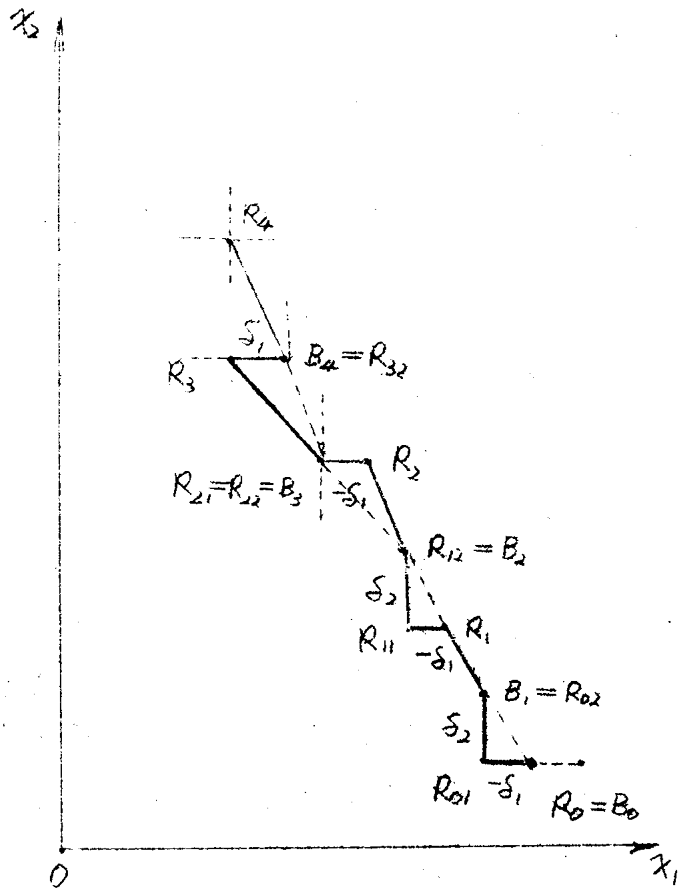
从上可以看出，步长加速法的搜索过程可以述之如下：从一个参考点出发施行步骤I得到一个新的基点，如果这个基点能确立，说明平移成功，并在这二个接续的基点的联线上进行再平移，这样又得一个新的参考点。如果平移不成功，新基点不能建立，探测作废，退回原基点并把这个基点看作新的初始点，按照上述规则重新开始新的探测过程。由于规定 $(x_1^0, x_2^0)==R_0=B_0$，故给出一个初始点后，按上述规则交替使用步骤I、II，循环往复，可最后找到最优点。在整个计算过程中，步长一般不改变，只有在初始点或某一个新的初始点上，施行步骤I以后失败时，才将步长缩小，开始新的计算过程。所需注意的是，平移以后要检验平移是否成功。

试举例说明这个方法的步骤（图1.4）

求某一目标函数 $f(x_1, x_2)$ 的极大点。给定步长 δ_1 与 δ_2，取 $B_0 = R_0$。从初始点 R_0 出发向右移动，若失败，返回 R_0，再向左移动，若成功，记 $R_{01} = R_0 - \delta_1 e_1$。至此，水平探测结束。自 R_{01} 向上移动成功，记此点为 R_{02}，$R_{02} = R_{01} + \delta_2 e_2$，因为 $f(R_{02}) > f(B_0)$，记 $B_1 = R_{02}$，至此从 R_0 出发的局部探测完成，新基点 B_1 确立。继之施行步骤Ⅱ，从 B_1 出发，作新参考点 $R_1 = 2B_1 - B_0$。从 B_1 平移至 R_1 后是否平移成功，需要从 R_1 出发施步骤Ⅰ加以检验。由于上次向左移动成功，故从 R_1 出发是向左移动，移动成功，记得的点为 R_{11}，$R_{11} = R_1 - \delta_1 e_1$，自 R_{11} 向上移动成功，记得的点为 R_{12}，$R_{12} = R_{11} + \delta_2 e_2$。由于有 $f(R_{12}) > f(B_1)$，故平移成功，确立一个新基点 B_2，$B_2 = R_{12}$，同时又找到一个新的上升方向 $B_2 - B_1$，按照上述规则，从 B_2 出发又可再进行平移至 R_2，

$$R_2 = B_2 + (B_2 - B_1) = 2B_2 - B_1$$

R_2 是一个新参考点，故从 R_2 出发施行步



骤I。从 R_2 出发向左移动成功，得 R_{21}，$R_{21}=R_2-\delta_1 e_1$，但自 R_{21} 向上向下移动均失败，记 $R_{22}=R_{21}$。由于 $f(R_{22})>f(B_2)$，加大步长向前移动得新参考点 $R_3=2B_3-B_2$。自 R_3 向左移动失败，向右移动成功，记 $R_{31}=R_3+\delta_1 e_1$，自 R_{31} 向上向下移动均失败，记 $R_{32}=R_{31}$。由于 $f(R_{32})>f(B_3)$，故得新基点 B_4 从 B_4 出发再施行步骤II，又得一个新参考点 $R_4=2B_4-B_3$，但从 R_4 向各个方向探测均失败，故 $R_{42}=R_4$，同时由于 $f(R_4)<f(B_4)$，因此平移失败，从 R_{42} 退回 B_4。按上述规则，从 B_4 开始新的探索，把 B_4 看成新的初始点，但自 B_4 施行步骤I失败，故缩短步长，继续探测过程。

§2.2 n维的标法

n维即 n 个因素时的标法与二个因素时的标法没有原则上的不同，由于是 n 个因素，此时有 n 个坐标轴方向 $e_1=(1,\underbrace{0,\cdots 0}_{n-1\text{个}})$，$e_2=(0,1,0,\cdots 0)$，……$e_n=(\underbrace{0,0,\cdots 0}_{n-1\text{个}},1)$，一个'点'有 n 个坐标分量，在一个点周围的局部探测按 e_1，e_2，…… e_n 的方向依次进行，探测 n 次。计标过程简述如下：

1、局部探测：

给定一初始点 $R_0=(x_1^0,x_2^0,\cdots x_n^0)$，也就是各因素取一固定值，令 $B_0=R_0$，也就是初始点既看作参考点又是基点。从 R_0 出发沿 e_1 的正方向移动，得点 $R_0+\delta_1 e_1=(x_1^0,x_2^0\cdots x_n^0)+$ $+\delta_1(1,0,\cdots 0)=(x_1^0+\delta_1,x_2^0,\cdots x_n^0)$，若 $f(R_0+\delta_1 e_1)>f(R_0)$，则