钻井布局的优化模型

制作者：　　重庆大学数理学院　　李传东

摘　要

针对勘探部门在钻井找矿时，如何进行最优钻井布局的问题，进行了深入的分析和讨论，利用一维搜索、二维搜索、三维搜索得到不同条件下最多可利用旧井数的算法。作者在分析了几篇优秀论文后，对本问题的认识有了很大的提高。现就该问题的基本概念、分析方法、模型建立及其求解等谈一下看法，供大家参考。

目　　录

问题重述

勘探部门在某地区找矿，初步勘探时期已零散的在若干位置上钻井，取得了地质资料，进入系统勘探时期后，要在一个区域内按纵横等距的网格点来布置井位，进行“撒网式”全面钻探。由于钻一口井的费用很高，如果新设计的井位与原有井位重合（或相当接近），便可利用旧井的地质资料，不必打这口新井。因此，应该尽量利用，少打新井，以节约钻探费用。比如钻一口新井的费用是500万元，利用旧井资料的费用为10万元，则利用一口旧井就节约费用490万元。

设平面上有n个点P_i，其坐标为（a_i，b_i），$i=1,2,3,\ldots,n$，表示已有的n个井位。新布置的井位是一个正方形网格的所有结点（所谓“正方形网格”是指每个格子都是正方形的网格；结点是指纵线和横线的交叉点）。假定每个格子的边长（井位的纵横间距）都是1单位（比如100米）。整个网格是可以在平面上任意移动的。若一个已知点P_i与某个网格结点的距离不超过给定误差（=0.05单位），则认为P_i处的旧井资料可以利用，不必在结点X_i处打新井。

为进行辅助决策，勘探部门要求我们研究以下问题：

1）假定网格的横向和纵向是固定的（比如东西向和南北向），并规定两点间的距离为横向距离（横坐标之差的绝对值）与纵向距离（纵坐标之差的绝对值）的最大值，在平面上平行移动网格N，使可利用的旧井数尽可能大。试提供数值计算方法，并对下面的数值例子用计算机进行计算。

2）在欧氏距离的误差意义下，考虑网格的横向和纵向不固定（可以旋转）的情形，给出算法并计算结果。

3）如果有n口旧井，给出判定这些旧井均可利用的条件和算法（你可以任意选择一种距离）。

数值例子 n=12个点的坐标如下表所示												
i	1	2	3	4	5	6	7	8	9	10	11	12
a_i	0.50	1.41	3.00	3.37	3.40	4.72	4.72	5.43	7.57	8.38	8.98	9.50
b_i	2.00	3.50	1.50	3.51	5.50	2.00	6.24	4.10	2.01	4.50	3.41	0.80

概念说明

解此题不需要专门的知识，方法是初等的，但一些概念要清楚，一些好的答卷都是由于把握住了这些概念；而大部分答卷中的错误都是源于这方面的疏忽。

a) 取整运算

研究网格点与其他点的关系必然要用取整运算。常用的有如下两种：

[x]=不大于x的最大整数（x的整数部分），

r(x)=[x+0.5](x按四舍五入规则取整)。

此外，还会用到上取整和下取整。这里[x]相当于下取整。

我们可以分别表示按下取整及四舍五入取整的小数部分为：

$$\{x\}= x - [x] , f(x)=|x - r(x)|$$

用这些记号来表示一个点与格点的距离是方便的。

b) 距离概念

本题考虑两种距离。给定两点**P(a,b)**及**X(x,y)**，第一种距离是所谓L_∞模距离：

$$d(P,X)=\max\{|x - a| , |y - b|\},$$

在平面上通常称为纵横距离。第二种距离是欧氏距离，即L_2模距离；

$$\rho(P,X)=\sqrt{(x-a)^2+(y-b)^2}$$

在平面上也称为直线距离。

大部分答卷都通过对坐标“去整运算”（去掉整数部分），将n个点P_i变换到单位正方形

$$Q=[0,1;0,1]=\{(x,y)|\ 0 \le x \le 1,\ 0 \le y \le 1\}$$

内，一个关键之处是上述变换使距离概念发生了变化。事实上，由于函数{x}的周期性，正方形Q的左边界与右边界是粘合的，上边界和下边界是粘合的。因此，严格的说，Q不再是一个平面区域，而是一个环面（像汽车轮胎一样的曲面），其中的距离发生了变化：在Q中两点 $\bar{P}_i(\bar{a}_i,\bar{b}_i)$ 及 $\bar{P}_j(\bar{a}_j,\bar{b}_j)$ 的纵横距离变为

$$\begin{aligned} d'(\bar{P}_i,\bar{P}_j) &= \max\{\min\{|\ \bar{a}_i-\bar{a}_j\ |,1-|\ \bar{a}_i-\bar{a}_j\ |\},\min\{|\ \bar{b}_i-\bar{b}_j\ |,1-|\ \bar{b}_i-\bar{b}_j\ |\}\} \\ &= \max\{f(|\ \bar{a}_i-\bar{a}_j\ |),f(|\ \bar{b}_i-\bar{b}_j\ |)\} \end{aligned}$$

注意：不强求学生的答卷中明确写出这种变化后的距离公式，但要意识到这种距离的变化。好的答卷都对此作了处理。例如，将正方形**Q**向外扩展一定程度，将粘合的边界摊开，然后运用原来平面上的距离；或者当要求 $d'(\bar{P}_i,\bar{P}_j)\le 2\varepsilon$ 时等价的写出

$$\begin{cases}|\bar{a}_i-\bar{a}_j|\le 2\varepsilon\text{或}1-|\bar{a}_i-\bar{a}_j|\le 2\varepsilon\\ |\bar{b}_i-\bar{b}_j|\le 2\varepsilon\text{或}1-|\bar{b}_i-\bar{b}_j|\le 2\varepsilon\end{cases}$$

而那些不考虑距离变化的答卷必然导出错误的结论。

c) 关于“同时被利用”

对集合上的一种性质来说，如集合A具有这种性质能推出A的任意子集也具有，数学上便称它为“独立性”。这种独立性十分普遍，例如：(a)一组向量线性无关；(b)图G中一组边的导出子图不含圈；(c)一组随机事件相互独立。在我们的问题中一组旧井可以被利用，则它的一部分亦然；所以“同时被利用”也是一种独立性。在这种意义上，我们的问题实际上是一类最大独立集问题。现在要问：一个集合中任意两个元素组成的集合是独立（两两独立）的能否推出整个集合是独立的？回答说否定的。对目前的问题来说，我们却有如下有趣的结论：

(i) 在坐标轴方向固定的情形下，对纵横距离来说，一组旧井可以同时被利用当且仅当其中任意两口井可以同时被利用。

(ii) 在坐标轴方向固定的情形下，对欧氏距离来说，一组旧井可以同时被利用当且仅当其中任意三口井可以同时被利用。

在坐标轴方向不固定的情形，难以得到此类结论。根据结论(i)，不少同学利用两口井同时被利用的条件建立一个二元关系，进而将问题转化为图论中的最大团问题，这是对的，而且对问题作了很好的刻画。但是，对欧氏距离，甚至坐标轴方向不固定的情形也这样做就错了。根据结论(ii)，用覆盖的方法解决问题就容易实现了。

分析与建模

1. 问题1）的分析与建模

思路一：对坐标作“去整运算”

根据题义，网格的方向是固定的，对于任意一点P，当网格纵横平移整数个单位时，P相对于最近大的网格结点的距离是不变的，即当P在网格上纵横平移整数个单位时，P相对于网格结点的距离不变。于是，我们把所有的旧井点都纵横平移整数个单位，使他们都落在同一个网格单元（单位正方形）内。但考虑到函数{x}的周期性，该正方形的左边界与右边界是粘合的，上边界和下边界是粘合的。将正方形向外扩展一定程度，将粘合的边界摊开，

这里可以摊开成以（a,b）、(a+1+ε，b+1+ε)为对角顶点的正方形，然后运用原来平面上的距离。要使尽可能多的网格结点与旧井点的距离不大于ε，等价于让尽可能多的旧井点落在以结点为中心的、以2ε为边长的正方形Q内。问题1）转化成移动Q使其盖住尽可能多的旧井点。

设旧井点P原来的坐标为（x,y），则按上述方法得到新坐标为（x′，y′）：

如果单位正方形以（0，0）和（1，1）为对角顶点：

$x'=x-[x]$，或者$x'=x-[x]+1$；$y'=y-[y]$，或者$y'=y-[y]+1$

当旧井点$P_i(x_i', y_i')$可利用时记$\mu_i=1$，否则记$\mu_i=0$。设正方形Q的中心坐标为(x, y)于是可得如下最优化问题：

$$\max \sum_{i=1}^{n} \mu_i$$

s.t.

$$\max\{|x-x_i'|, |y-y_i'|\} \le \varepsilon$$

思路二：直接法

不妨设$a_i,b_i \geq 0(i=1,2,\ldots,n)$，即所有井点$P_i$处在原坐标系Oxy的第一象限。在网格N上取第一象限中离原点O最近的结点为(s,t)，其中$0 \leq s<1$，$0 \leq t<1$。这个结点看作是网格N上一个参照点或新坐标系的原点，它可以在上述单位正方形Q内移动。于是N的任意结点可以表示为$(s+p,t+q)$，其中$p,q \in Z$（这里Z表示整数集）。

对问题1）而言，当且仅当ε-邻域

$$U_\varepsilon(P_i)=\{(x,y)\,|\,a_i-\varepsilon\le x\le a_i+\varepsilon, b_i-\varepsilon\le y\le b_i+\varepsilon\}$$

中存在结点，即不等式组

$$\begin{cases} a_i-\varepsilon\le s+x_i\le a_i+\varepsilon \\ b_i-\varepsilon\le t+y_i\le b_i+\varepsilon \\ 0\le s,t<1 \\ x_i,y_i\in Z \end{cases} \qquad (3.1)$$

有解时，P_i是可利用的。对给定的(s,t)，（3.1）有解的充要条件是

$$\begin{cases} a_i-\varepsilon-s\le [a_i+\varepsilon-s] \\ b_i-\varepsilon-t\le [b_i+\varepsilon-t] \end{cases} \qquad (3.2)$$

P_i可利用时记μ_i=1,否则μ_i=0。这样一来，问题1）归结为如下的最优化问题：

$$
\begin{cases}
\max \sum_{i=1}^{n} \mu_i \\
\text{s.t.} \quad (a_i - \varepsilon - s - [a_i + \varepsilon - s]\mu_i \le 0, i = 1,2,\cdots,n \\
\qquad (b_i - \varepsilon - t - [b_i + \varepsilon - t]\mu_i \le 0, i = 1,2,\cdots,n \\
\qquad 0 \le s, t \le 1 \\
\qquad \mu_i \in \{0,1\}, i = 1,2,\cdots,n
\end{cases}
$$

这是一个非线性的混合整数规划。

2． 问题2）的分析与建模

问题2）也可以写成数学规划的形式，但较复杂，我们换一种写法。为简单起见，假设坐标系经旋转φ（$0\le\varphi\le\pi/2$）之后，点P_{i}-的坐标仍记为（a_i,b_i）。所以现在的a_i和b_i是φ的函数（坐标变换式从略）。在网格N上，处于第一象限内而离原点最近的结点仍记为(s,t)。欧氏距离的ε-领域为：

$$S_\varepsilon(P_i)=\{(x,y)|(x-a)^2+(y-b)^2\le\varepsilon^2\}。$$

显然，$S_\varepsilon(P_i)\subset U_\varepsilon(P_i)$。当(3.2)成立时，含于$U_\varepsilon(P_i)$中唯一的结点是

$$（s+[a_i+\varepsilon-s]，t+[b_i+\varepsilon-t]）。$$

因此，P_i是可利用的当且仅当

$$\begin{cases} (a_i-\varepsilon-s\le[a_i+\varepsilon-s] \\ b_i-\varepsilon-t\le[b_i+\varepsilon-t] \\ (s+[a_i+\varepsilon-s]-a_i)^2+(t+[b_i+\varepsilon-t]-b_i)^2\le\varepsilon^2 \end{cases} \qquad (3.3)$$

这样一来，问题2）是如下的最优化问题：

$$\sum_{i=1}^{n}\mu_i(s,t)$$

其中

$$\mu_i(s,t)=\begin{cases} 1, & \text{当}(3.3)\text{成立} \\ 0, & \text{否则} \end{cases}$$

$$0\le s,t\le 1$$

对问题3）的分析

关于问题3），可以采用下面的两种途径：代数途径、几何途径。

- 代数途径
- 几何途径

代数途径

首先考虑纵横距离，并假设坐标轴方向（转角φ）已经给定，**P_i** 的坐标为**(a_i,b_{-i})**。所谓代数途径，就是研究不等式组的整数解存在性问题。根据前一节的数学模型，**n**口井同时被利用的充要条件是对**i=1,2,…,.n**的以下两个不等式组有解：

$$\begin{cases} a_i-\varepsilon\le s+x_i\le a_i+\varepsilon(i=1,2,\cdots,n) \\ x_i\in Z, \qquad 0\le s<1 \end{cases} \tag{4.1}$$

$$\begin{cases} b_i-\varepsilon\le t+y_i\le b_i+\varepsilon(i=1,2,\cdots,n) \\ y_i\in Z, \qquad 0\le t<1 \end{cases} \tag{4.2}$$

他们的形式是一样的，只要研究其中的一组，比如(4.1)。由于

$$a_i - \varepsilon - 1 < a_i - \varepsilon - s \le x_i \le a_i + \varepsilon - s \le a_i + \varepsilon$$

所以

$$[a_i - \varepsilon] \le x_i \le [a_i + \varepsilon]$$

注意：左边的不等式是由$a_i - \varepsilon - 1 < x_i$及$x_i$为整数得到。记 $\alpha_i = [a_i - \varepsilon]$，$\beta_i = [a_i + \varepsilon]$。根据题目的具体意义，可以约定 $\varepsilon < 1/4$。由此可知$0 \le \beta_i - \alpha_i \le 1$。所以（4.1）的整变量$x_i$只有两个可能值$\alpha_i$和$\beta_i$。

命题　当且仅当如下两条件之一成立时，不等式组(4.1)有解：

$$\text{(i)} \quad \max_{1\le i\le n}(a_i - \varepsilon - \alpha_i) \le \operatorname{mix}_{1\le i\le n}(a_i + \varepsilon - \alpha_i)$$

$$\text{(ii)} \quad \max_{1\le i\le n}(a_i - \varepsilon - \beta_i) \le \operatorname{mix}_{1\le i\le n}(a_i + \varepsilon - \beta_i)$$

定理1　设坐标轴方向给定，对纵横距离而言，n口旧井可同时被利用的充分必要条件是

$$\max_{1\le i\le n}(a_i-\varepsilon-[a_i-\varepsilon])\le\min_{i\le i\le n}(a_i+\varepsilon-[a_i-\varepsilon])\text{ 或}$$

$$\max_{1\le i\le n}(a_i-\varepsilon-[a_i+\varepsilon])\le\min_{i\le i\le n}(a_i+\varepsilon-[a_i+\varepsilon])$$

$$\max_{1\le i\le n}(b_i-\varepsilon-[b_i-\varepsilon])\le\min_{i\le i\le n}(b_i+\varepsilon-[b_i-\varepsilon])\text{ 或}$$

$$\max_{1\le i\le n}(b_i-\varepsilon-[b_i+\varepsilon])\le\min_{i\le i\le n}(b_i+\varepsilon-[b_i+\varepsilon])$$

注意：计算复杂性为O(n)。这是因为计算每一个max和min的比较次数均为n – 1。

这种途径，对欧氏距离及旋转角度φ变动的情形，未能得到简洁的判定条件。

几何途径

首先假设坐标轴方向是固定的，并考虑纵横距离。将点$P_i(a_i, b_i)$变换为

$$\overline{P}_i(\overline{a}_i, \overline{b}_i),\ \text{其中}\ \overline{a}_i = \{a_i\} = a_i - [a_i],\ \overline{b}_i = \{b_i\} = b_i - [b_i],\ i = 1, 2, \cdots, n$$

这就使变换后的点落入单位正方形Q=[0,1;0,1]之中。同时，一个由参照点(s,t)确定的网格的所有结点都变换到Q中的同一个点(s,t)。对这样的变换来说，点Pi与其最接近的结点的距离等于$\overline{P}_i$与(s,t)的距离。因此，$P_1,P_2,\ldots,P_n$可同时被利用的充分必要条件是坐标变换后的点同时被利用。

定理2 设坐标轴方向给定，对纵横距离而言，n口旧井可同时被利用的充要条件是： 对任意的i,j(1≤s,t ≤n)均有

$$\begin{cases} |\bar{a}_i - \bar{a}_j| \le 2\varepsilon \text{或} 1- |\bar{a}_i - \bar{a}_j| \le 2\varepsilon \\ |\bar{b}_i - \bar{b}_j| \le 2\varepsilon \text{或} 1- |\bar{b}_i - \bar{b}_j| \le 2\varepsilon \end{cases}$$

其次，讨论欧氏距离，并假定坐标轴方向（转角φ）已经给定。如前，可利用去整运算，将所有点变换到正方形Q内。注意到环面Q的周期性，Q中两点的直线距离也发生了变化（特别是接近Q的边界的点）。为避免这种周期性的影响，许多答卷都将Q的边界向外扩展，并将接近边界的点变为两个或三个点（新点称为镜像点）。这样，两点间的距离就变成了一点及其镜像点与零一点及其镜像点之间的最短距离。下面的“切补法”也是可行的：

1. 将$Q=[0,1;0,1]$分成三部分：

$$A=[0,1/4;0,1],B=[1/4,3/4;0,1],C=[3/4,1;0,1]$$

2. 若A,C内均有井点，则B内不可能有井点（否则，B内井点到$A\cup C$内井点的最大距离$>1/2>2\varepsilon$）,他们不可能同时被利用，这种情况可以排除）。于是可以把区域A切补到[1，1+1/4；0，1]的位置，得到正方形Q'。

3. 同理对Q'作纵向的处理，得到正方形Q''。这样就可以在Q''中直接运用欧氏距离了。

我们仍不妨假设n个井点分布于单位正方形Q内，从几何上说，n个井点同时被利用就是存在一个半径为ε的圆域覆盖住这n个点，其圆心就是网格N的参照点(s,t)。下面是一个贫穷按段条件。

定理3　平面上n个点可以被一个半径为ε圆域覆盖当切仅当如下两条件成立：

1)　任意两点距离≤2ε；

2)　任意形成锐角三角形的三点的外界圆半径≤ε。

推论　设坐标轴方向给定，对欧氏距离而言，n个井点可同时被利用的充要条件是任意三点可同时被利用。

进一步研究的问题

1) 对问题3），可以考虑坐标轴方向不定的情形，较困难。

2) 可以考虑更多的距离概念，如常用的L_1距离

3) 在实际问题中，如果网格的格子不是单位正方形，而是遍常有一定限制的矩形。