

雷达对抗讲义

目 录

第 1 章 基础知识	1
1.1 符号规定	1
1.2 矩阵微积分	1
1.3 复矩阵微积分	4
1.4 常见统计量	6
1.5 拉格朗日乘数法	6
第 2 章 无源定位	9
2.1 测向定位	9
2.1.1 二维平面	9
2.1.2 三维空间	13
2.2 时差定位	16
2.3 频差定位	18
第 3 章 雷达信号处理基础	21
3.1 雷达信号的基本形式	21
3.2 目标参数测量	24
3.2.1 测距	24
3.2.2 测角	26
3.3 测速	28
第 4 章 超分辨率参数测量	29
4.1 谱估计	29
4.1.1 补零法	29
4.1.2 Capon 算法	30
4.1.3 MUSIC 算法	30
4.2 来波方位估计	30
4.2.1 Capon 算法	30
4.2.2 MUSIC 算法	31
4.2.3 ESPRIT 算法	32

4.3 时延估计	32
4.3.1 互相关法	32
4.3.2 相位拟合法	33
 第 5 章 雷达抗干扰	 35
5.1 最小二乘	35
5.2 独立成分分析	36
5.3 主偏度分析	38
5.3.1 偏度的定义	38
5.3.2 数据白化	39
5.3.3 张量基本运算	40
5.3.4 统计量映射图	42
5.3.5 任意方向的偏度	44
5.3.6 协偏度张量的计算	45
5.3.7 模型与求解	47

第 2 章 无源定位

雷达是一种主动探测设备，通过发射电磁波，并接收目标反射的回波来确定目标的位置和速度。但在真实战场中，这是一种很危险的行为，极易暴露自身位置，从而遭到敌方的打击。因此，无源定位技术应运而生。无源定位是指通过被动接收目标发射的电磁波信号，来获取目标的相关信息。常见的无源定位技术有测向定位、时差定位和频差定位等。无源定位技术的优点在于隐蔽性强，难以被敌方发现，因此在现代战争中得到了广泛应用。

2.1 测向定位

2.1.1 二维平面

利用机械扫描，或者电子扫描等方式，雷达基站可以获得所接收的电子的方位信息。本节以二维平面为例，显然，一个雷达基站确定了一条过目标以及雷达基站自身的直线。如图 2.1 所示，在二维平面上，至少二个雷达基站可以确定目标位置。

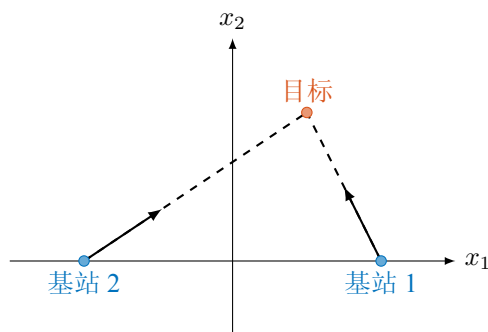


图 2.1 测向定位示意图

但在实际应用中，由于会受到噪声和干扰的影响，雷达基站获得的目标角度信息往往存在误差，为了提高定位精度，通常需要多个基站进行联合定位。但此时一个新的问题出现了，由于误差的存在，多个基站所确定的直线并不一定相交于同一点（如图 2.2 所示），此时目标的位置该如何确定呢？

设有 N 个基站，且第 n 个基站确定的过目标的直线方程为 $a_{n1}x_1 + a_{n2}x_2 = b_n$,

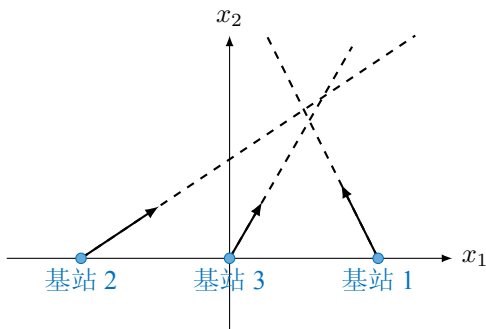


图 2.2 无源定位可能存在观测误差

则 N 个基站确定了如下的方程组：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \\ \vdots \\ a_{N1}x_1 + a_{N2}x_2 = b_N \end{cases}.$$

由于存在误差，以上方程组并不一定有解。但我们希望最终求解得到的目标位置 (x_1, x_2) 尽可能满足所有直线方程。此时，新的问题出现了，我们该如何衡量解满足方程的程度呢？在测向定位中，这个指标很容易就可以得到了，即目标到某个基站所确定的直线的距离。根据点到直线的距离公式，目标到第 n 个基站所确定的直线的距离为：

$$d_n = \frac{|a_{n1}x + a_{n2}y - b_n|}{\sqrt{a_{n1}^2 + a_{n2}^2}}.$$

因此，可以得到如下的目标函数：

$$\min_{x_1, x_2} \sum_{n=1}^N d_n^2 = \sum_{n=1}^N \frac{(a_{n1}x + a_{n2}y - b_n)^2}{a_{n1}^2 + a_{n2}^2}.$$

注意到，比起最小化距离和，我们更倾向于最小化距离的平方和，这样可以避免绝对值带来的分段函数。

此外，目标到第 n 个基站所确定的直线的距离表达式中存在一个带根号的分母，这不便于计算；而对于一条直线方程，等式两边同除一个非零的数并不会改变这条直线。因此，我们可以通过归一化的方式，将分母变为 1，从而简化计算。具体地，我

们可以将第 n 个基站所确定的直线方程两边同除以 $\sqrt{a_{n1}^2 + a_{n2}^2}$ ，从而得到新的方程组：

$$\begin{cases} \frac{a_{11}}{\sqrt{a_{11}^2 + a_{12}^2}}x + \frac{a_{12}}{\sqrt{a_{11}^2 + a_{12}^2}}y = \frac{b_1}{\sqrt{a_{11}^2 + a_{12}^2}} \\ \frac{a_{21}}{\sqrt{a_{21}^2 + a_{22}^2}}x + \frac{a_{22}}{\sqrt{a_{21}^2 + a_{22}^2}}y = \frac{b_2}{\sqrt{a_{21}^2 + a_{22}^2}} \\ \vdots \\ \frac{a_{N1}}{\sqrt{a_{N1}^2 + a_{N2}^2}}x + \frac{a_{N2}}{\sqrt{a_{N1}^2 + a_{N2}^2}}y = \frac{b_N}{\sqrt{a_{N1}^2 + a_{N2}^2}} \end{cases}.$$

记

$$\begin{cases} \hat{a}_{n1} = \frac{a_{n1}}{\sqrt{a_{n1}^2 + a_{n2}^2}} \\ \hat{a}_{n2} = \frac{a_{n2}}{\sqrt{a_{n1}^2 + a_{n2}^2}} \\ \hat{b}_n = \frac{b_n}{\sqrt{a_{n1}^2 + a_{n2}^2}} \end{cases},$$

则目标到第 n 个基站所确定的直线的距离可以简化为：

$$d_n = |\hat{a}_{n1}x + \hat{a}_{n2}y - \hat{b}_n|.$$

为了简化计算，我们可以将方程组写成矩阵的形式：

$$\begin{bmatrix} \hat{a}_{11} & \hat{a}_{12} \\ \hat{a}_{21} & \hat{a}_{22} \\ \vdots & \vdots \\ \hat{a}_{N1} & \hat{a}_{N2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \vdots \\ \hat{b}_N \end{bmatrix}$$

也即 $\mathbf{Ax} = \mathbf{b}$ 。此时目标到所有基站确定的直线的距离构成的向量 $\mathbf{d} = [d_1 \ d_2 \ \dots \ d_N]^T$ 有如下的表达式：

$$d_n = |(\mathbf{Ax} - \mathbf{b})_n|.$$

而目标函数则可以简化为：

$$\min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{b}\|^2.$$

对于该目标函数，我们有如下定理：

定理 2.1 $f(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|^2$ 为一个凸函数。

证明 根据定义，函数 $f(\mathbf{x})$ 为凸函数当且仅当对于任意的 $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ ，都有

$$\frac{f(\mathbf{x}_1) + f(\mathbf{x}_2)}{2} \geq f\left(\frac{\mathbf{x}_1 + \mathbf{x}_2}{2}\right).$$

将目标函数 $f(\mathbf{x})$ 展开，有

$$f(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|^2 = (\mathbf{Ax} - \mathbf{b})^T (\mathbf{Ax} - \mathbf{b}) = \mathbf{x}^T \mathbf{A}^T \mathbf{Ax} - 2\mathbf{b}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{b}.$$

因此, $f\left(\frac{\mathbf{x}_1+\mathbf{x}_2}{2}\right)$ 可以展开为:

$$f\left(\frac{\mathbf{x}_1+\mathbf{x}_2}{2}\right) = \frac{1}{4}\mathbf{x}_1^T\mathbf{A}^T\mathbf{A}\mathbf{x}_1 + \frac{1}{4}\mathbf{x}_2^T\mathbf{A}^T\mathbf{A}\mathbf{x}_2 + \frac{1}{2}\mathbf{x}_1^T\mathbf{A}^T\mathbf{A}\mathbf{x}_2 - \mathbf{b}^T\mathbf{A}(\mathbf{x}_1+\mathbf{x}_2) + \mathbf{b}^T\mathbf{b}.$$

与此同时, $\frac{f(\mathbf{x}_1)+f(\mathbf{x}_2)}{2}$ 可以展开为:

$$\frac{f(\mathbf{x}_1)+f(\mathbf{x}_2)}{2} = \frac{1}{2}\mathbf{x}_1^T\mathbf{A}^T\mathbf{A}\mathbf{x}_1 + \frac{1}{2}\mathbf{x}_2^T\mathbf{A}^T\mathbf{A}\mathbf{x}_2 - \mathbf{b}^T\mathbf{A}(\mathbf{x}_1+\mathbf{x}_2) + \mathbf{b}^T\mathbf{b}.$$

所以, 只需证明

$$\begin{aligned}\frac{f(\mathbf{x}_1)+f(\mathbf{x}_2)}{2} - f\left(\frac{\mathbf{x}_1+\mathbf{x}_2}{2}\right) &= \frac{1}{4}\mathbf{x}_1^T\mathbf{A}^T\mathbf{A}\mathbf{x}_1 + \frac{1}{4}\mathbf{x}_2^T\mathbf{A}^T\mathbf{A}\mathbf{x}_2 - \frac{1}{2}\mathbf{x}_1^T\mathbf{A}^T\mathbf{A}\mathbf{x}_2 \\ &= \frac{1}{4}(\mathbf{x}_1-\mathbf{x}_2)^T\mathbf{A}^T\mathbf{A}(\mathbf{x}_1-\mathbf{x}_2) \geq 0.\end{aligned}$$

记 $\mathbf{v} = \mathbf{A}(\mathbf{x}_1 - \mathbf{x}_2)$, 显然有

$$\mathbf{v}^T\mathbf{v} = (\mathbf{x}_1 - \mathbf{x}_2)^T\mathbf{A}^T\mathbf{A}(\mathbf{x}_1 - \mathbf{x}_2) \geq 0.$$

因此, $f(\mathbf{x})$ 为一个凸函数。 ■

对于一个凸函数, 我们可以通过求解其导数为 0 的点来找到最优解。具体而言, 对于目标函数 $f(\mathbf{x})$, 我们有

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = 2\mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) = 0.$$

因此, 最优解为

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}.$$

这就是著名的最小二乘解。

例 2.1 设二维平面上有三个基站, 确定了如下的方程组:

$$\begin{cases} 2x + y = 4 \\ 2x - 3y = -4 \\ 1.9x - 1.1y = 0 \end{cases}.$$

请求解目标位置 (x, y) 。

解 首先将直线方程归一化, 得到

$$\begin{cases} \frac{2}{\sqrt{5}}x + \frac{1}{\sqrt{5}}y = \frac{4}{\sqrt{5}} \\ \frac{2}{\sqrt{13}}x - \frac{3}{\sqrt{13}}y = -\frac{4}{\sqrt{13}} \\ \frac{1.9}{\sqrt{4.82}}x - \frac{1.1}{\sqrt{4.82}}y = 0 \end{cases}.$$

构建对应的矩阵和向量：

$$\mathbf{A} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{13}} & -\frac{3}{\sqrt{13}} \\ \frac{1.9}{\sqrt{4.82}} & -\frac{1.1}{\sqrt{4.82}} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \frac{4}{\sqrt{5}} \\ -\frac{4}{\sqrt{13}} \\ 0 \end{bmatrix}.$$

代入最小二乘解公式，得到

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} = \begin{bmatrix} 1.0540 \\ 1.9635 \end{bmatrix}.$$

从图 2.3 可以看到，最小二乘得到的解正位于目标函数的最小值处。这意味该点到所有基站确定的直线的距离的平方和最小，也就是目标最有可能出现的位置。

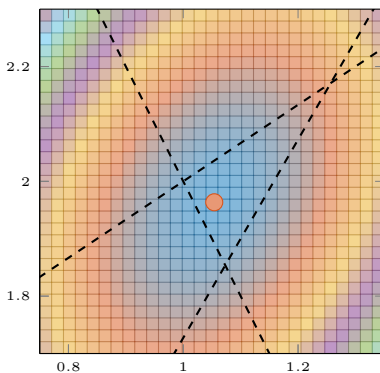


图 2.3 测向定位目标函数

2.1.2 三维空间

在实际应用中，我们更多地需要确定目标在三维空间中的位置，此时又该如何求解呢？注意到，在三维空间中，雷达基站并不是直接确定一条过目标的直线，而是分别确定目标所在的方位角和俯仰角，也就是确定了两个过目标的平面。因此，类似于二维平面中的情况，我们可以通过至少两个基站确定的四个平面来求解目标的位置。

设共确定了 N 个平面，对应的方程组为：

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ \vdots \\ a_{N1}x + a_{N2}y + a_{N3}z = b_N \end{cases}.$$

同样，我们可以将方程组归一化，得到

$$\begin{cases} \frac{a_{11}}{\sqrt{a_{11}^2+a_{12}^2+a_{13}^2}}x + \frac{a_{12}}{\sqrt{a_{11}^2+a_{12}^2+a_{13}^2}}y + \frac{a_{13}}{\sqrt{a_{11}^2+a_{12}^2+a_{13}^2}}z = \frac{b_1}{\sqrt{a_{11}^2+a_{12}^2+a_{13}^2}} \\ \frac{a_{21}}{\sqrt{a_{21}^2+a_{22}^2+a_{23}^2}}x + \frac{a_{22}}{\sqrt{a_{21}^2+a_{22}^2+a_{23}^2}}y + \frac{a_{23}}{\sqrt{a_{21}^2+a_{22}^2+a_{23}^2}}z = \frac{b_2}{\sqrt{a_{21}^2+a_{22}^2+a_{23}^2}} \\ \vdots \\ \frac{a_{N1}}{\sqrt{a_{N1}^2+a_{N2}^2+a_{N3}^2}}x + \frac{a_{N2}}{\sqrt{a_{N1}^2+a_{N2}^2+a_{N3}^2}}y + \frac{a_{N3}}{\sqrt{a_{N1}^2+a_{N2}^2+a_{N3}^2}}z = \frac{b_N}{\sqrt{a_{N1}^2+a_{N2}^2+a_{N3}^2}} \end{cases}$$

记

$$\begin{cases} \hat{a}_{n1} = \frac{a_{n1}}{\sqrt{a_{n1}^2+a_{n2}^2+a_{n3}^2}} \\ \hat{a}_{n2} = \frac{a_{n2}}{\sqrt{a_{n1}^2+a_{n2}^2+a_{n3}^2}} \\ \hat{a}_{n3} = \frac{a_{n3}}{\sqrt{a_{n1}^2+a_{n2}^2+a_{n3}^2}} \\ \hat{b}_n = \frac{b_n}{\sqrt{a_{n1}^2+a_{n2}^2+a_{n3}^2}} \end{cases},$$

则目标到第 n 个基站所确定的平面的距离可以简化为：

$$d_n = |\hat{a}_{n1}x + \hat{a}_{n2}y + \hat{a}_{n3}z - \hat{b}_n|.$$

后面的求解过程与二维平面中的情况相同，这里不再赘述。

当然，我们也可以直接使用三维空间中的直线方程来求解目标位置。设有 N 个基站，且第 n 个基站确定的过目标的直线方程参数方程为

$$\begin{cases} x_1 = l_{n1} + k_{n1}t \\ x_2 = l_{n2} + k_{n2}t \\ x_3 = l_{n3} + k_{n3}t \end{cases}.$$

记

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{l}_n = \begin{bmatrix} l_{n1} \\ l_{n2} \\ l_{n3} \end{bmatrix}, \quad \mathbf{k}_n = \begin{bmatrix} k_{n1} \\ k_{n2} \\ k_{n3} \end{bmatrix},$$

则第 n 个基站确定的直线方程可以简化为

$$\mathbf{x} = \mathbf{l}_n + \mathbf{k}_n t.$$

注意到， $\mathbf{k}_n$ 代表了直线的方向，因此不妨令其模长为 1，即 $\mathbf{k}_n^T \mathbf{k}_n = 1$ 。

设目标的位置为 $\mathbf{w}$ ，则目标到第 n 个基站所确定的直线的距离平方为

$$d_n^2 = \min_t \|\mathbf{w} - \mathbf{l}_n - \mathbf{k}_n t\|^2.$$

因此，我们可以得到如下的目标函数：

$$\min_{\mathbf{w}} \sum_{n=1}^N d_n^2.$$

首先，让我们来求解目标到第 n 个基站所确定的直线的距离平方。首先将对应的目标函数展开：

$$\begin{aligned} d_n^2 &= \min_t \|\mathbf{w} - \mathbf{l}_n - \mathbf{k}_n t\|^2 \\ &= \min_t (\mathbf{w} - \mathbf{l}_n - \mathbf{k}_n t)^\top (\mathbf{w} - \mathbf{l}_n - \mathbf{k}_n t) \\ &= \min_t t^2 (\mathbf{k}_n^\top \mathbf{k}_n) - 2t(\mathbf{w} - \mathbf{l}_n)^\top \mathbf{k}_n + (\mathbf{w} - \mathbf{l}_n)^\top (\mathbf{w} - \mathbf{l}_n). \end{aligned}$$

该函数为一个关于 t 的二次函数，显然其最小值出现在

$$t = \frac{(\mathbf{w} - \mathbf{l}_n)^\top \mathbf{k}_n}{\mathbf{k}_n^\top \mathbf{k}_n} = (\mathbf{w} - \mathbf{l}_n)^\top \mathbf{k}_n.$$

将其代入到距离的平方公式中，得到

$$\begin{aligned} d_n^2 &= \|\mathbf{w} - \mathbf{l}_n - \mathbf{k}_n (\mathbf{w} - \mathbf{l}_n)^\top \mathbf{k}_n\|^2 \\ &= \|\mathbf{w} - \mathbf{l}_n - \mathbf{k}_n \mathbf{k}_n^\top (\mathbf{w} - \mathbf{l}_n)\|^2 \\ &= \left\| \left(\mathbf{I} - \mathbf{k}_n \mathbf{k}_n^\top \right) (\mathbf{w} - \mathbf{l}_n) \right\|^2 \\ &= (\mathbf{w} - \mathbf{l}_n)^\top \left(\mathbf{I} - \mathbf{k}_n \mathbf{k}_n^\top \right)^\top \left(\mathbf{I} - \mathbf{k}_n \mathbf{k}_n^\top \right) (\mathbf{w} - \mathbf{l}_n) \\ &= (\mathbf{w} - \mathbf{l}_n)^\top \left(\mathbf{I} - \mathbf{k}_n \mathbf{k}_n^\top \right) (\mathbf{w} - \mathbf{l}_n) \end{aligned}$$

记 $\mathbf{P}_n = \mathbf{I} - \mathbf{k}_n \mathbf{k}_n^\top$ 因此，目标函数可以简化为

$$\begin{aligned} \min_{\mathbf{w}} \sum_{n=1}^N d_n^2 &= \min_{\mathbf{w}} \sum_{n=1}^N (\mathbf{w} - \mathbf{l}_n)^\top \mathbf{P}_n (\mathbf{w} - \mathbf{l}_n) \\ &= \min_{\mathbf{w}} \mathbf{w}^\top \left(\sum_{n=1}^N \mathbf{P}_n \right) \mathbf{w} - 2 \sum_{n=1}^N \left(\mathbf{l}_n^\top \mathbf{P}_n \right) \mathbf{w} + \sum_{n=1}^N \mathbf{l}_n^\top \mathbf{P}_n \mathbf{l}_n. \end{aligned}$$

记

$$\mathbf{P} = \sum_{n=1}^N \mathbf{P}_n, \quad \mathbf{l} = \left(\sum_{n=1}^N \mathbf{l}_n^\top \mathbf{P}_n \right)^\top,$$

则目标函数等价于

$$\min_{\mathbf{w}} \mathbf{w}^\top \mathbf{P} \mathbf{w} - 2\mathbf{l}^\top \mathbf{w}.$$

其关于 $\mathbf{w}$ 的导数为

$$\frac{\partial f(\mathbf{w})}{\partial \mathbf{w}} = 2\mathbf{P}\mathbf{w} - 2\mathbf{l}.$$

因此, 最优解为

$$\mathbf{w} = \mathbf{P}^{-1}\mathbf{l}.$$

2.2 时差定位

利用目标发射的电磁波抵达不同位置的基站的时间不同, 也可以确定目标的位置。设有 2 个基站, 其中基站 1 和基站 2 分别于 t_1 和 t_2 时刻接收到目标发射的电磁波信号, 那么目标到两个基站之间的距离差为

$$d = c(t_2 - t_1),$$

其中 c 为光速。根据几何知识, 我们知道与两个固定的点 (称为焦点) 的距离差是常数的点的轨迹为双曲线, 如图 2.4 所示。

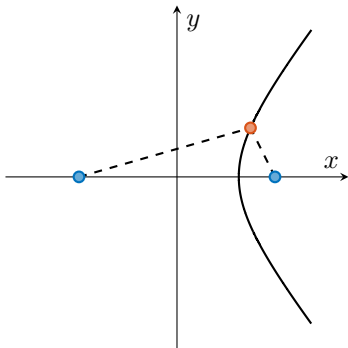


图 2.4 双曲线示例

显然, 只有两个基站时, 只能确定目标所在的双曲线。不过, 可以通过增加一个基站, 通过求解两个双曲线的交点¹来确定目标的位置, 如图 2.5 所示。

实际应用中, 时差的测量同样也会受到噪声和干扰的影响, 因此我们可以通过增加基站的数量来提高定位精度。此时, 多个基站确定的双曲线并不一定交于一点。设有 N 个基站, 且第 i 个基站接收到目标发射的电磁波信号的时刻为 t_i , 则目标到第 i 个基站和第 j 个基站的距离差为

$$d_{ij} = c(t_j - t_i).$$

设目标位置为 $\mathbf{x} = [x_1 \ x_2]^T$, 第 i 个基站的位置为 $\mathbf{l}_i = [l_{1i} \ l_{2i}]^T$, 则目标到第 i 个基站的距离为

$$d_i = \|\mathbf{x} - \mathbf{l}_i\| = \sqrt{(\mathbf{x} - \mathbf{l}_i)^T(\mathbf{x} - \mathbf{l}_i)} = \sqrt{(x_1 - l_{1i})^2 + (x_2 - l_{2i})^2}.$$

¹实际上是三个双曲线, 并且这三个双曲线必然交于一点

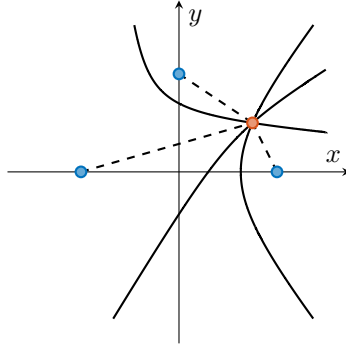


图 2.5 时差定位示例

因此，我们可以构建如下的目标函数：

$$\min_{\mathbf{x}} \sum_{i=1}^N \sum_{j=1}^N (d_i - d_j - d_{ij})^2.$$

很容易发现，该目标函数并不是一个凸函数，对于这样的目标函数，我们可以使用梯度下降法来求解最优解。

设目标函数为 $f(\mathbf{x})$ ，其关于 $\mathbf{x}$ 的导数为

$$\begin{aligned} \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} &= 2 \sum_{i=1}^N \sum_{j=1}^N (d_i - d_j - d_{ij}) \left(\frac{\partial d_i}{\partial \mathbf{x}} - \frac{\partial d_j}{\partial \mathbf{x}} - \frac{\partial d_{ij}}{\partial \mathbf{x}} \right) \\ &= \sum_{i=1}^N \sum_{j=1}^N (d_i - d_j - d_{ij}) \left(\frac{\mathbf{x} - \mathbf{l}_i}{d_i} - \frac{\mathbf{x} - \mathbf{l}_j}{d_j} \right) \\ &= \sum_{i=1}^N \sum_{j=1}^N (d_i - d_j - d_{ij}) (\mathbf{x}_i - \mathbf{x}_j), \end{aligned}$$

其中 $\mathbf{x}_i = \frac{\mathbf{x} - \mathbf{l}_i}{d_i}$ 。因此，我们可以得到如下的迭代公式：

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \eta \frac{\partial f(\mathbf{x}_k)}{\partial \mathbf{x}}.$$

其中 η 为学习率。

在实践中，学习率的选择非常重要，过大的学习率会导致目标函数震荡，甚至发散；而过小的学习率则会导致收敛速度过慢。为了避免学习率选择不当导致的震荡和发散，我们可以使用牛顿法来求解最优解。牛顿法的基本思想是通过二阶导数来加速

收敛速度。对于目标函数为 $f(\mathbf{x})$ ，其关于 $\mathbf{x}$ 的二阶导数，也即海森（Hessian）矩阵为：

$$\begin{aligned} \frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x}^2} &= \frac{\left(\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \right)^T}{\partial \mathbf{x}} \\ &= \sum_{i=1}^N \sum_{j=1}^N (\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j) + (d_i - d_j - d_{ij}) \left(\frac{\mathbf{I} - \mathbf{x}_i \mathbf{x}_i^T}{d_i} - \frac{\mathbf{I} - \mathbf{x}_j \mathbf{x}_j^T}{d_j} \right). \end{aligned}$$

记 $\mathbf{d}_\mathbf{x} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}$ ， $\mathbf{H}_\mathbf{x} = \frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x}^2}$ ，则牛顿法的迭代公式为

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{H}_{\mathbf{x}_k}^{-1} \mathbf{d}_{\mathbf{x}_k}.$$

从图 2.6 可以看到，不同的学习率下，梯度下降法的收敛速度差异很大，并且越大的学习率并不一定收敛得更快。由于牛顿法的收敛速度为二次收敛，而梯度下降法的收敛速度为线性收敛，因此牛顿法的收敛速度远快于梯度下降法。

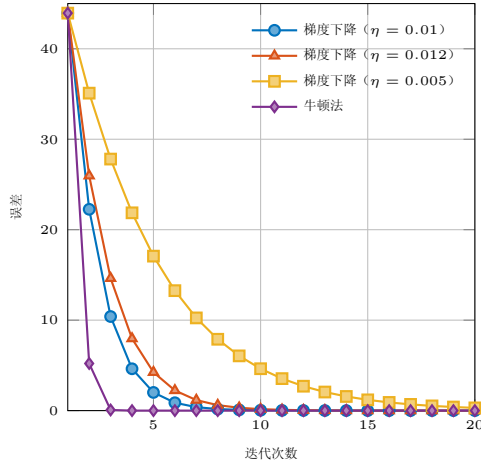


图 2.6 梯度下降法和牛顿法收敛速度对比

此外，需要注意的是，时差定位的目标函数可能会存在多个局部最优解。为了避免陷入局部最优解，有很多方法可以使用，例如随机初始化、模拟退火等。这里我们不再赘述。

2.3 频差定位

利用多普勒效应，我们可以通过运动中的接收基站来确定静止目标的位置。设有目标位置为 $\mathbf{x}$ ，并且在 i 个时刻，接收基站的位置为 $\mathbf{l}_i$ ，速度为 $\mathbf{v}_i$ 。如目标发射的信号频率为 f_0 ，则根据多普勒效应，接收基站在第 i 个时刻接收到的信号频率为

$$f_i = f_0 \left(1 + \frac{\mathbf{v}_i^T (\mathbf{x} - \mathbf{l}_i)}{c \|\mathbf{x} - \mathbf{l}_i\|} \right).$$

但是由于目标发射的信号频率 f_0 是未知的，因此我们无法直接使用上述公式来确定目标的位置。为此，我们需要再次进行观测，设在第 j 个时刻，接收基站的位置为 $\mathbf{l}_j$ ，速度为 $\mathbf{v}_j$ ，则接收基站在第 j 个时刻接收到的信号频率为

$$f_j = f_0 \left(1 + \frac{\mathbf{v}_j^T (\mathbf{x} - \mathbf{l}_j)}{c \|\mathbf{x} - \mathbf{l}_j\|} \right).$$

将两式相比，我们可以得到

$$\frac{f_i}{f_j} = \frac{1 + \frac{\mathbf{v}_i^T (\mathbf{x} - \mathbf{l}_i)}{c \|\mathbf{x} - \mathbf{l}_i\|}}{1 + \frac{\mathbf{v}_j^T (\mathbf{x} - \mathbf{l}_j)}{c \|\mathbf{x} - \mathbf{l}_j\|}}.$$

对其进行变形，得到

$$f_j \left(1 + \frac{\mathbf{v}_i^T (\mathbf{x} - \mathbf{l}_i)}{c \|\mathbf{x} - \mathbf{l}_i\|} \right) - f_i \left(1 + \frac{\mathbf{v}_j^T (\mathbf{x} - \mathbf{l}_j)}{c \|\mathbf{x} - \mathbf{l}_j\|} \right) = 0.$$

记

$$f_{ij}(\mathbf{x}) = f_j \left(1 + \frac{\mathbf{v}_i^T (\mathbf{x} - \mathbf{l}_i)}{c \|\mathbf{x} - \mathbf{l}_i\|} \right) - f_i \left(1 + \frac{\mathbf{v}_j^T (\mathbf{x} - \mathbf{l}_j)}{c \|\mathbf{x} - \mathbf{l}_j\|} \right),$$

则函数 $|f_{ij}(\mathbf{x})|$ 具有图 2.7 所示的等高线。

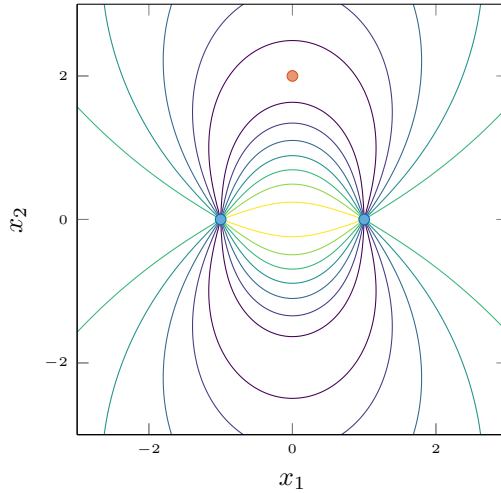


图 2.7 频差定位误差等高线（红色点为目标，蓝色点为观测点，基站自左向右匀速移动）

同样地，我们可以通过增加基站的数量来提高定位精度，设共进行了 N 次观测，对应的频差定位目标函数为

$$\min_{\mathbf{x}} \sum_{i=1}^N \sum_{j=1}^N f_{ij}^2(\mathbf{x}).$$

该目标函数的求解与时差定位类似，这里不再赘述。