

无线信道参数估计算法

对通信系统分析之前，首先需要建立一个能够完整反映该系统的**传输信道的模型**，其模型中的参数包括接收端接收信号的**时延**，发射端和接收端的**联合方向分布**等，获取这些参数的过程就是无线信道参数估计的过程。

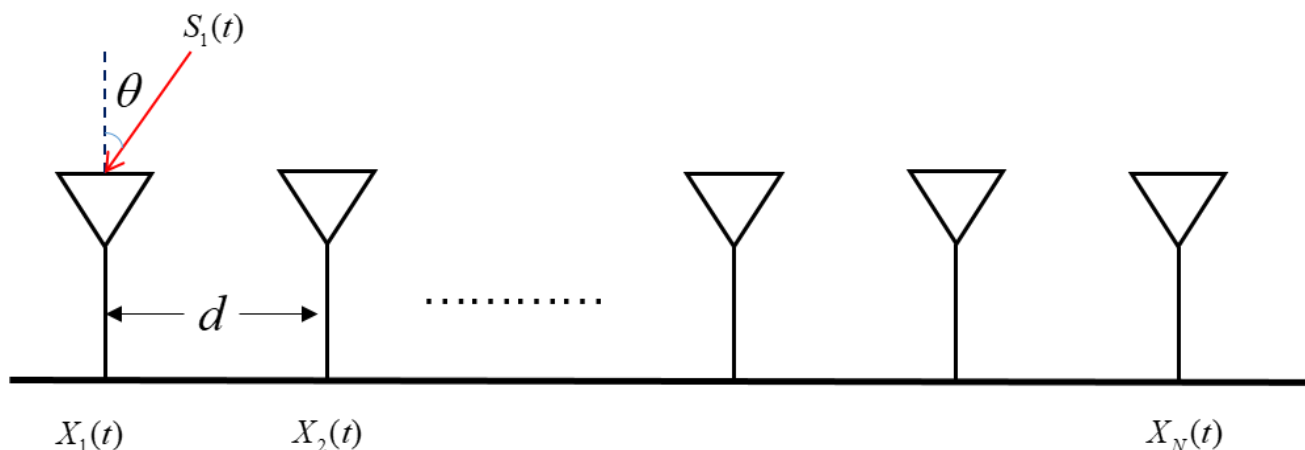
方法可以大致分为三类：

- 谱估计算法：多重信号分类法
- 参数子空间估计算法：旋转不变技术估计信号参数算法
- 确定性参数估计算法：期望最大算法以及演进的空间交替广义期望最大算法

1. 多重信号分类法

MUSIC 算法原理

MUSIC 算法，叫做多信号分类算法 (Multiple Signal classification)。MUSIC 算法的基本思想是将任意阵列输出数据的协方差矩阵进行特征分解，从而得到与信号分量相对应的信号子空间和信号分量相正交的噪声子空间，然后利用这两个子空间的正交性来估计信号的参数 (入射方向、极化信息和信号强度)。MUSIC 算法是一种高分辨率的、高精度的无线信道参数估计算法，而且该算法对于天线阵列的形状没有特殊要求，具有普遍的适用性，只要已知天线阵的布阵形式，无论是直线阵列还是圆阵列，不管阵元是否是等间距分布，都可以得到高分辨率的结果。但是该算法对于入射信号的要求非常高，应用该算法的前提是**入射信号必须是互不相干的**。



考虑由 N 个阵元组成的均匀线性阵列 (Uniform Linear Array)，阵列之间的间距为半波长 $d = \lambda/2$ ，假设有 M 个远场信号源，它们的功率相同且互不相干，分别从不同的方向入射到接收阵列，其中入射信号和噪声互不相关。以最左侧的阵元为参考点，各个阵元的位置相比前一个增加距离 d 。信号入射方向角为 θ ，表示入射信号与线阵法线的夹角。一般假设信号源数 M 小于阵元数目 N 。

在第 k 次快拍，得到的接收数据向量为：

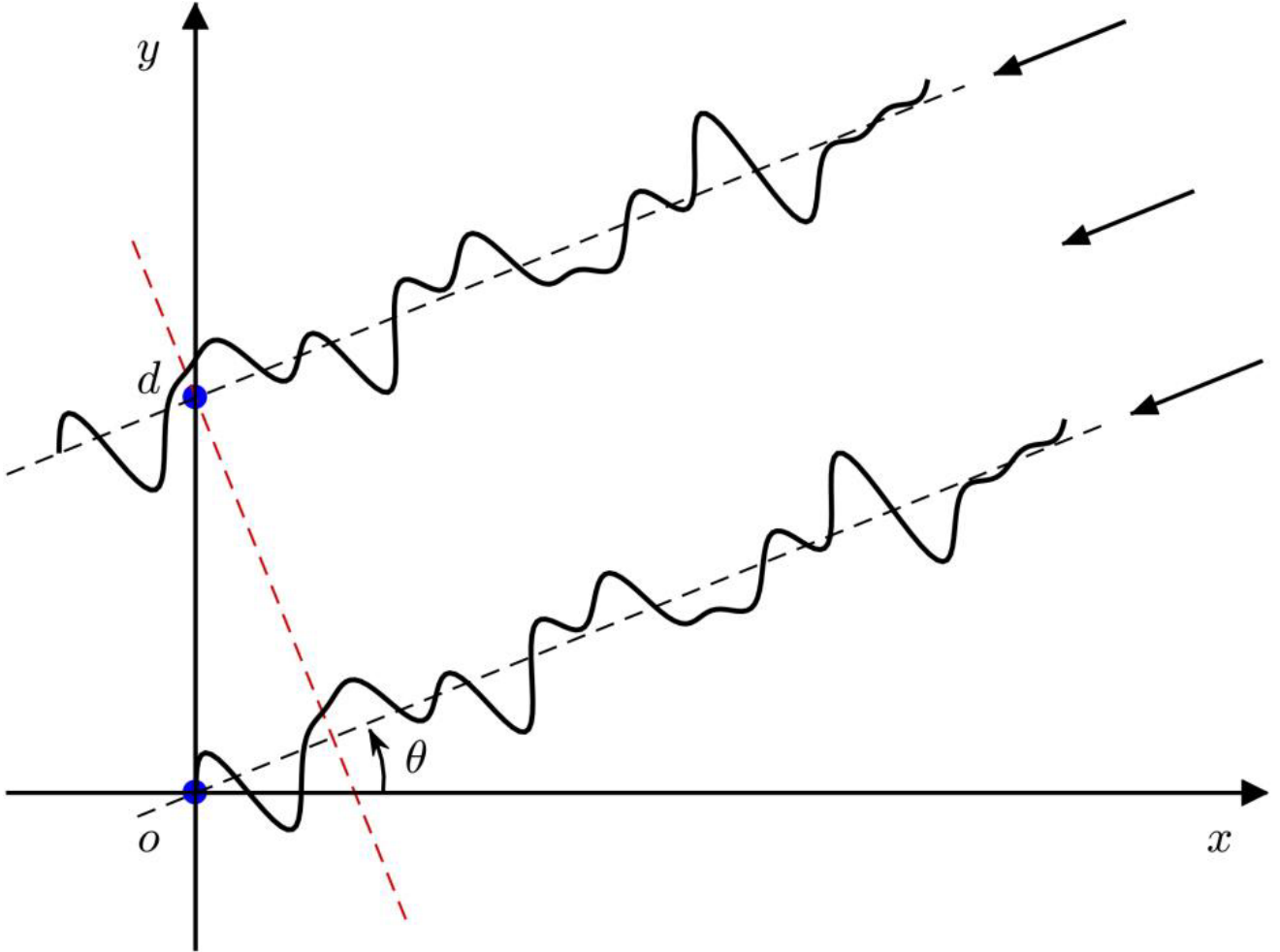
$$X(k) = A(\theta)S(k) + N(k), k = 1, 2, \dots, K$$

其中, $A(\theta) = [a(\theta_1), a(\theta_2), \dots, a(\theta_M)]$ 是 $N \times M$ 的阵列矩阵, 又称为阵列流形向量或者方向响应向量。 $S(k) = [s_1(k), s_2(k), \dots, s_M(k)]^T$ 是 M 个入射信号矢量。 $X(k) = [x_1(k), x_2(k), \dots, x_N(k)]^T$ 是 N 个输出信号矢量。 $N(k) = [n_1(k), n_2(k), \dots, n_N(k)]^T$ 是阵列噪声矢量。一般假设噪声为加性高斯即: $\mathbb{E}(n_i(k)) = 0$, $\mathbb{E}(n_i(k)n_i(k)^H) = \sigma^2 I$ 。

阵列矩阵的元素满足:

$$a(\theta_i) = [1, e^{-j2\pi d \sin \theta_i / \lambda}, \dots, e^{-j(N-1)2\pi d \sin \theta_i / \lambda}]$$

其中, λ 是中心波长, θ_i 为信号源的入射角。因此, 问题转化为对 X 进行多点采样, 通过对采样的输出信号 $X(k)$ 估计出信号源的波达方向角 θ_i 。



由上图中两束方向确定的信号发射到两个不同阵元, 在同一时刻采样的信号值为随机值, 待估的方向一定。于是, 输出信号的协方差矩阵可以表示为:

$$\begin{aligned} R_{XX} &= \mathbb{E}[X(k)X(k)^H] \\ &= \mathbb{E}[A(\theta)S(k)S(k)^H A(\theta)^H] + \mathbb{E}[N(k)N(k)^H] \\ &= A(\theta)\mathbb{E}[S(k)S(k)^H] A(\theta)^H + R_N \\ &= A R_S A^H + \sigma^2 I \end{aligned}$$

其中 R_S 是输入信号的协方差矩阵。由于信号和噪声相互独立, 数据协方差可分解为信号子空间和噪声子空间。由于 R_{XX} 是对称矩阵, 然后我们对其进行特征值分解, 得到

$$R_{XX} = U \Sigma U^H = U_S \Sigma_S U_S^H + U_N \Sigma_N U_N^H$$

其中 $U = [U_S, U_N]$, $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_N^2)$, U_S 是由大的特征值对应的特征向量张成的子空间, 即信号子空间。 U_N 是由小的特征值对应的特征矢量张成的子空间, 即为噪声子空间。

$$\Sigma = \begin{bmatrix} \Sigma_S & \\ & \Sigma_N \end{bmatrix} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{bmatrix},$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M > \lambda_{M+1} = \dots = \lambda_N = \sigma^2$,

$$\Sigma_S = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_M \end{bmatrix},$$

$$\Sigma_N = \begin{bmatrix} \lambda_{M+1} & & & \\ & \lambda_{M+2} & & \\ & & \ddots & \\ & & & \lambda_N \end{bmatrix}.$$

由于信号子空间和噪声子空间是相互正交的，有

$$\begin{aligned} R_{XX}U_N &= [U_S, U_N] \Sigma \begin{bmatrix} U_S^H \\ U_N^H \end{bmatrix} U_N \\ &= [U_S, U_N] \Sigma \begin{bmatrix} \mathbf{0} \\ I \end{bmatrix} \\ &= [U_S, U_N] \begin{bmatrix} \sigma_S^2 & \\ & \sigma_N^2 \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ I \end{bmatrix} \\ &= \sigma_N^2 U_N \end{aligned}$$

结合 $R_{XX}U_N = AR_SA^H U_N + \sigma^2 I U_N = \sigma_N^2 U_N$ 成立，可得

$$\boxed{AR_SA^H U_N = \mathbf{0}}$$

$$U_N^H AR_SA^H U_N = (A^H U_N)^H R_SA^H U_N = \mathbf{0}$$

因为 $A \neq 0$, $R_S \neq \mathbf{0}$ ，可推出上式的充分必要条件： $A^H U_N = \mathbf{0}$ 。这个性质也可以从另外一方面说明，已知模型为 $X = AS + N$ ，在无噪声情况下的信号 $x(t) = \sum_{i=1}^M s_i(t)a_i(\theta) \rightarrow x(t) \in \text{span}\{a_1, a_2, \dots, a_M\}$ 。**协方差矩阵大特征值对应的特征向量张成的空间与入射信号的导向矢量张成的空间是同一个空间，即**

$$U_S = [e_1, e_2, \dots, e_M], \quad U_N = [e_{M+1}, e_{M+2}, \dots, e_N],$$

$$\text{span}\{e_1, e_2, \dots, e_M\} = \text{span}\{a_1, a_2, \dots, a_M\}.$$

信号子空间和噪声子空间是相互正交，可知信号子空间的导向矢量 $A(\theta)$ 和噪声子空间也正交，即

$$\mathbf{a}^H(\theta)U_N = \mathbf{0}, \quad \theta = \{\theta_1, \theta_2, \dots, \theta_M\}.$$

然而实际采集的接收数据矩阵有限长，首先获取输入信号的大量采样值 $X = \{x_1, x_2, \dots, x_K\}$ ，进而可以估计输入信号的协方差矩阵，达到极大似然估计的目的：

$$\hat{R}_x = \frac{1}{K} \sum_{k=1}^K x_k x_k^H$$

其中， K 是快拍数，然后对 \hat{R}_x 特征值分解： $\hat{R}_x = \hat{U}\Sigma\hat{U}^H$ 。同理得到两个子空间：

$$\begin{aligned} \hat{R}_x &= \hat{U}\Sigma\hat{U}^H \\ &= \hat{U}_S \Sigma_S \hat{U}_S^H + \hat{U}_N \Sigma_N \hat{U}_N^H \end{aligned}$$

由于实际的干扰因素， $\mathbf{a}^H(\theta)U_N = \mathbf{0}$ 不完全成立，即不会完全满足正交性。这里采用最小优化搜索来估计波达方向：

$$\theta_{MUSIC} = \arg \min_{\theta} \mathbf{a}^H(\theta)\hat{U}_N \hat{U}_N^H \mathbf{a}(\theta)$$

从上式可以看出，由于阵列方向向量与噪声子空间正交的缘故，信号入射方向上会出现极小值。通过上式整理得到空间谱函数，进行极大值的谱峰搜索，可得

$$P_{MUSIC}(\theta) = \frac{1}{\mathbf{a}^H(\theta) \hat{\mathbf{U}}_N \hat{\mathbf{U}}_N^H \mathbf{a}(\theta)}$$

MUSIC 实验仿真

```

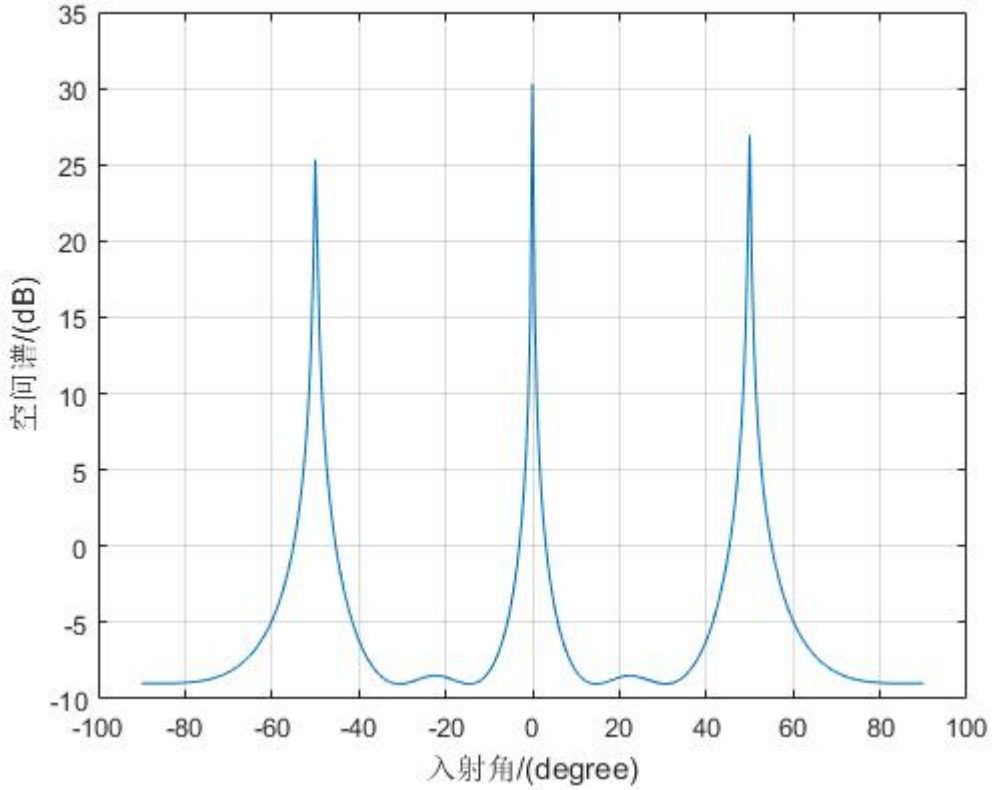
clc; clear; close all;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% MUSIC %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
derad = pi/180;
N = 8;           % 阵元个数
M = 3;           % 信源数目
theta = [-50 0 50]; % 待估计角度
snr = 10;        % 信噪比
K = 1024;        % 快拍数 10/1024

dd = 0.5;        % 阵元间距 d=lamda/2
d=0:dd:(N-1)*dd;
A=exp(-1i*2*pi*d.*sin(theta*derad));

S=randn(M,K); X=A*S;
X1=awgn(X,snr,'measured');
Rxx=X1*X1'/K;
%InvS=inv(Rxx);
[EV,D]=eig(Rxx); % 特征向量 特征值
EVA=diag(D)';
[EVA,I]=sort(EVA); % 从小到大排列 返回索引
%EVA=flip1r(EVA); % 反转元素
EV=flip1r(EV(:,I)); % 按列反转特征向量

for iang = 1:361 % 遍历
    angle(iang)=(iang-181)/2;
    phim=derad*angle(iang);
    a=exp(-1i*2*pi*d.*sin(phim)).';
    L=M;
    En=EV(:,L+1:N);
    %SP(iang)=(a'*a)/(a'*En*En'*a);
    SP(iang)=1/(a'*En*En'*a);
end
SP=abs(SP);
%SPmax=max(SP);
%SP=10*log10(SP/SPmax); % 归一化
SP=10*log10(SP);
h=plot(angle,SP);
set(h,'Linewidth',0.5);
xlabel('入射角/(degree)');
ylabel('空间谱/(dB)');
%axis([-100 100 -40 60]);
set(gca, 'XTick',[-100:20:100]);
grid on;

```



ROOT-MUSIC 算法

基本的算法中，结果的精度与扫描的精度直接相关，增大扫描精度也即要减小扫描间隔，也即用运算量来换取精度。求根算法是对 MUSIC 基本算法的演进，降低运算量。在求根算法中可以用多项式求根的方法来替代算法中的谱峰搜索。考虑以下模型：

$$X(k) = A(\theta)S(k) + \xi(k), k = 1, 2, \dots, K$$

在基本算法中，数据协方差矩阵是可以分解为相互正交的信号子空间和噪声子空间。阵列矩阵投影到噪声子空间，并在角度功率谱上进行谱峰搜索。现定义 $z = \exp(j2\pi d \sin \theta / \lambda)$ ，则阵列矩阵转化为：

$$\begin{aligned} \mathbf{a}(\theta_i) &= [1, e^{-j2\pi d \sin \theta_i / \lambda}, \dots, e^{-j(N-1)2\pi d \sin \theta_i / \lambda}]^T \\ &= [1, e^{j2\pi d \sin \theta_i / \lambda}, \dots, e^{j(N-1)2\pi d \sin \theta_i / \lambda}]^H \\ &= [1, z(\theta_i), \dots, z^{N-1}(\theta_i)]^H \end{aligned}$$

结合子空间 $U_N = [e_{M+1}, e_{M+2}, \dots, e_N]$ ，极大值的谱峰搜索函数可以改写为：

$$\begin{aligned} P_{MUSIC}^{-1}(\theta) = f(z) &= \mathbf{a}^H(\theta) \mathbf{U}_N \mathbf{U}_N^H \mathbf{a}(\theta) \\ &= [1, z(\theta_i), \dots, z^{N-1}(\theta_i)] [e_{M+1}, e_{M+2}, \dots, e_N] \\ &\quad [e_{M+1}, e_{M+2}, \dots, e_N]^H [1, z(\theta_i), \dots, z^{N-1}(\theta_i)]^H \\ &= \mathbf{p}^H(z) \mathbf{U}_N \mathbf{U}_N^H \mathbf{p}(z) \end{aligned}$$

令 $f(z) = 0$ ，求得该多项式的根即为到达角的信息。但是该多项式存在共轭项 z^* ，因此修正上式为：

$$\begin{aligned}
f(z) &= \mathbf{p}^H(z) \mathbf{U}_N \mathbf{U}_N^H \mathbf{p}(z) \\
&= z^{N-1} \mathbf{p}^T(z^{-1}) \mathbf{U}_N \mathbf{U}_N^H \mathbf{p}(z) \\
&= \mathbf{q}(z) \mathbf{U}_N \mathbf{U}_N^H \mathbf{p}(z), \\
\text{where } \mathbf{q}(z) &= [z^{N-1}, z^{N-2}, \dots, 1].
\end{aligned}$$

基于以上变换，推导 $f(z)$ 的展开式：

$$\begin{aligned}
f(z) &= \mathbf{q}(z) \mathbf{U}_N \mathbf{U}_N^H \mathbf{p}(z), \\
&= \mathbf{q} (e_{M+1} e_{M+1}^*, e_{M+2} e_{M+2}^*, \dots, e_N e_N^*) \mathbf{p}(z) \\
&= [z^{N-1}, z^{N-2}, \dots, 1] \begin{bmatrix} b_{11} & & \\ & \ddots & \\ & & b_{NN} \end{bmatrix} \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{N-1} \end{bmatrix} \\
&= [(z^{N-1} b_{11} + z^{N-2} b_{21} + \dots + b_{N1}), \dots, \\
&\quad (z^{N-1} b_{1N} + z^{N-2} b_{2N} + \dots + b_{NN})] \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{N-1} \end{bmatrix} \\
&= (z^{N-1} b_{11} + z^{N-2} b_{21} + \dots + b_{N1}) + \dots \\
&\quad + (z^{N-1} b_{1N} + z^{N-2} b_{2N} + \dots + b_{NN}) z^{N-1} \\
&= b_{N1} + (b_{N-1,1} + b_{N2})z + \dots + b_{1N} z^{2(N-1)}
\end{aligned}$$

$f(z)$ 是阶数为 $2(N-1)$ 次幂的多项式，因此根据**代数基本定理**：任何一个非零的一元 n 次复系数多项式，都正好有 n 个复数根。也就是说 $f(z)$ 有 $(N-1)$ 对共轭根，根据定义 $z = \exp(j2\pi d \sin \theta / \lambda)$ 可知，当根满足定义时， $\mathbf{p}(z)$ 为导向量，且根正好在单位圆上（**模为1**），其中有 M 个根正好在单位圆上。因此，对于接近于单位圆上的根进行求解：

$$\begin{aligned}
\theta_i &= \arcsin\left(\frac{\lambda}{2\pi d} \arg\{\hat{z}\}\right), \quad i = 1, 2, \dots, M. \\
\text{where } \arg\{\exp(jw)\} &= w.
\end{aligned}$$

仿真代码如下：

```

%ROOT_MUSIC ALOGRITHM %DOA ESTIMATION BY ROOT_MUSIC
clc; clear; close all;
K=2; %信源数
M=8; %阵元数
L=200; %信号长度
w=[pi/4 pi/6].'; %信号频率
lamda=((2*pi*3e8)/w(1)+(2*pi*3e8)/w(2))/2; %信号波长
d_lamda=0.5; %阵元间距
snr=20; %信噪比
theta1=[45,60]; %信号入射角
for k=1:K
    A(:,k)=exp(-1j*[0:M-1]'*d_lamda*2*pi*sin(theta1(k)*pi/180)); %阵列流型
end
for kk=1:L
    s(:,kk)=sqrt(10.^((snr/2)/10))*exp(1j*w*(kk-1)); %仿真信号
end
x=A*s+(1/sqrt(2))*(randn(M,L)+1j*randn(M,L)); %加入高斯白噪声
R=(x*x')/L; %协方差矩阵
%%第一种方法
[V,D]=eig(R); %对方差矩阵进行特征分解
Un=V(:,1:M-K); %取噪声子空间
Gn=Un*Un';

```

```

a = zeros(2*M-1,1)'; %找出多项式的系数，并按阶数从高至低排列
for i=-(M-1):(M-1)
    a(i+M) = sum(diag(Gn,i));
end
a1=roots(a); %使用ROOTS函数求出多项式的根

a2=a1(abs(a1)<1); %找出在单位圆里且最接近单位圆的N个根
[lamda,I]=sort(abs(a2)-1)); %挑选出最接近单位圆的N个根
f=a2(I(1:K)); %计算信号到达方向角
source_doa=[asin(angle(f(1))/pi)*180/pi,asin(angle(f(2))/pi)*180/pi];
source_doa=sort(source_doa);
disp('source_doa');
disp(source_doa);
%%%%%%第二种方法%%%%%%%%%%
% [V,D]=eig(R);
% Un=V(:,1:M-K); % (4.5.7)
% Un1=Un(1:M,:);
% Un2=Un(K+1:M,:);
% T=[1 0 0 0 0]'; % (阵元数-信源数) *1
% c=Un1*inv(Un2)*T; % (K*(M-K))*((M-K)*(M-K))*((M-K)*1)=K*1
% c=[1,c(2,1),c(1,1)]; % (4.5.8)
% f=roots(c);
% source_doa=[asin(angle(f(1))/pi)*180/pi asin(angle(f(2))/pi)*180/pi];
% source_doa=sort(source_doa);
% disp('source_doa');
% disp(source_doa);

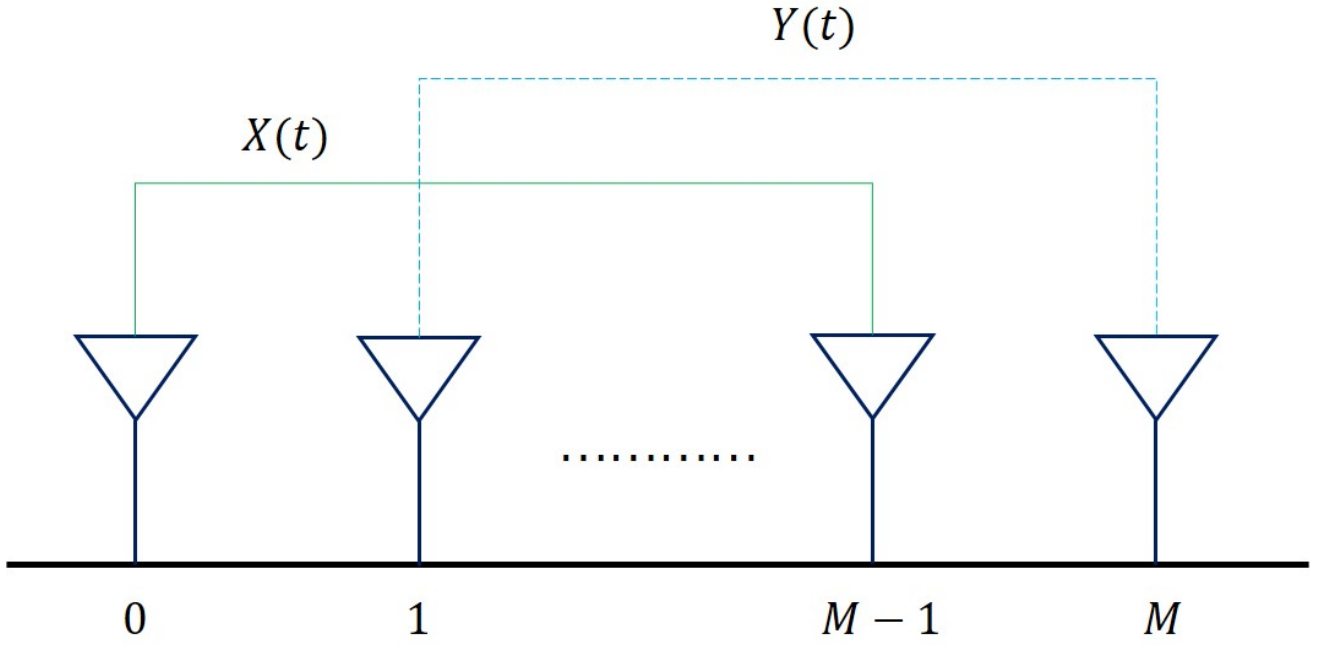
```

2. 旋转不变性子空间法

ESPRIT 算法原理

基于子空间的 DOA 估计算法包括两大类：MUSIC 类算法及 ESPRIT 类算法。旋转不变性子空间（ESPRIT）类算法和前面讲述的 MUSIC 类算法一样也需要对阵列接收数据协方差矩阵进行特征分解。但两者也有明显的不同点：其中 MUSIC 类算法是利用接收数据协方差矩阵的噪声子空间的正交性，而 ESPRIT 类算法是利用接收数据协方差矩阵的信号子空间的旋转不变性。

与 MUSIC 算法相比，ESPRIT 类算法优点在于计算量小，不需要在空间中不断的进行搜索谱峰。与 ROOT-MUSIC 算法相比，ESPRIT 类算法是利用各子阵的信号子空间之间的旋转不变性求出信号的入射角信息，而 ROOT-MUSIC 算法则是利用导向矢量与噪声子空间的正交性构造多项式并通过对多项式的求解来实现估计。



考虑一个有 M 个阵元的平面阵列结构，其元素被恒定位移矢量 Δ 分离。该阵列由两个完全相同的子阵构成，阵元数目均为 $(M-1)$ 个。其中子阵 1 由前 $(M-1)$ 个阵元组成，子阵 2 由后 $(M-1)$ 个阵元组成。假设子阵 1 向 x 轴正方向移动 Δ 时与子阵 2 重合。

假设位于远场的具有相同波长的 d ($d \leq (M-1)$) 个统计独立的窄带信号 $s_i(t)$, $i = 1, 2, \dots, d$ ，从方位方向 θ_i ($i = 1, 2, \dots, d$) 入射到阵列。分别用 X, Y 表示子阵 1, 2 的接收的来波信号矩阵。子阵 1, 2 的输出噪声为 N_1, N_2 ，假设它们为零均值且方差为 σ^2 的统计独立的高斯白噪声，且与信号矩阵不相关。

则对子阵 1、2 所接收的来波信号矩阵为：

$$\begin{aligned}
 X &= A_1 S + N_1 \\
 &= A S + N_1 \\
 &= [a(\theta_1), \dots, a(\theta_d)] [s_1(t), \dots, s_d(t)]^T + N_1 \\
 &= [x_1(t), x_2(t), \dots, x_{M-1}(t)]^T \\
 Y &= A_2 S + N_2 \\
 &= A \Phi S + N_2 \\
 &= [a(\theta_1), \dots, a(\theta_d)] \text{diag}[e^{j\phi_1}, \dots, e^{j\phi_d}] S + N_2 \\
 &= [a(\theta_1)e^{j\phi_1}, \dots, a(\theta_d)e^{j\phi_d}] S + N_2 \\
 &= [y_1(t), y_2(t), \dots, y_{M-1}(t)]^T
 \end{aligned}$$

其中， $a(\theta_i) = [a_1(\theta_i), \dots, a_{M-1}(\theta_i)]^T$ ， Φ 为两阵列间的延迟相位， $\phi_i = \frac{2\pi\Delta}{\lambda} \sin \theta_i$ 。由上式可知， Y 是由 X 旋转得到的。因此可以联立两个子阵的模型得到：

$$\begin{aligned}
 Z &= \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} A S + N_1 \\ A \Phi S + N_2 \end{bmatrix} \\
 &= \begin{bmatrix} A \\ A \Phi \end{bmatrix} S + \begin{bmatrix} N_1 \\ N_2 \end{bmatrix} \\
 &= \bar{A} S + N
 \end{aligned}$$

这里可得到协方差矩阵

$$\begin{aligned}
 R &= \mathbb{E}(Z Z^H) \\
 &= \mathbb{E}(\bar{A} S S^H \bar{A}^H + \bar{A} S N^H + N S^H \bar{A}^H + N N^H) \\
 &= \mathbb{E}(\bar{A} S S^H \bar{A}^H + N N^H) \\
 &= \bar{A} \mathbb{E}(S S^H) \bar{A}^H + \mathbb{E}(N N^H) \\
 &= \bar{A} R_S \bar{A}^H + R_N = \bar{A} R_S \bar{A}^H + \sigma^2 I
 \end{aligned}$$

取 K 个时刻的快拍数据，得：

$$Z(k) = \bar{A}(\theta)S(k) + N(k), \quad k = 1, 2, \dots, K$$

假设各到达角互不相同，即矩阵 \bar{A} 的列向量之间线性独立（**列满秩**）。实际中，只能利用输出采样值估计 R ，对协方差矩阵做特征分解：

$$\begin{aligned}\hat{R} &= \frac{1}{K} \sum_{t=1}^K Z(t)Z(t)^H, \\ \hat{R} &= U_S \Lambda_S U_S^H + U_N \Lambda_N U_N^H.\end{aligned}$$

其中， Z 的规模是 $2(M-1) \times 1$ ，因此 \hat{R} 是 $2(M-1) \times 2(M-1)$ 的， U_S 是 $2(M-1) \times d$ 的。由之前的性质知道 \bar{A} 与 U_S 张成的子空间一致，则一定存在唯一的非奇异矩阵 T ，使得 $U_S = \bar{A}T$ 。则可以分解为

$$U_S = \begin{bmatrix} U_X \\ U_Y \end{bmatrix} = \begin{bmatrix} A \\ A\Phi \end{bmatrix} T$$

从上式可立即得到 $U_Y = A\Phi T = U_X T^{-1} \Phi T = U_X \Psi$ ，其中 $\Psi = T^{-1} \Phi T$ 。一旦求出矩阵 Ψ ，就可以得到**旋转矩阵**：

$$\begin{aligned}\Phi &= T\Psi T^{-1} = \text{diag}[e^{j\phi_1}, \dots, e^{j\phi_d}], \\ \phi_i &= \frac{2\pi\Delta}{\lambda} \sin \theta_i \implies \hat{\theta}_i = \arcsin\left(\phi_i \times \frac{\lambda}{2\pi\Delta}\right).\end{aligned}$$

入射角度就可以通过上面的关系求得。从以上推导可知有如下关键信息：

$$\begin{aligned}\begin{bmatrix} U_X \\ U_Y \end{bmatrix} &= \begin{bmatrix} AT \\ A\Phi T \end{bmatrix} \implies \\ \text{span}\{A\} &= \text{span}\{U_X\} = \text{span}\{U_Y\}, \\ U_Y &= U_X \Psi \implies \Psi = U_X^\dagger U_Y.\end{aligned}$$

但实际应用中，我们只能得到有限的有噪声的测量数据去估计 $\text{span}\{U_S\}$ ，从而 $\text{span}\{U_S\} \neq \text{span}\{\bar{A}\}$ ，甚至 $\text{span}\{U_X\} \neq \text{span}\{U_Y\}$ 。所以这时我们无法找到 Ψ 使得 $U_Y = U_X \Psi$ 。必须寻找一个准则以获得 Ψ 的合适估计，通常情况下应用的是最小二乘 (LS) 估计。

TLS-ESPRIT 步骤

假设模型 $AX = B$ 中， A 已知且误差由 B 引起，若要获得 X 的估计通常应用标准的 LS 准则。这里容易证明 A 和 B 中同样存在噪声，那么 LS 准则只考虑了 U_Y 的噪声扰动，即 $U_Y + \Delta U_Y = U_X \Psi$ 就是不合适的。

考虑 A 和 B 都有噪声存在的准则是：整体最小二乘法 (TLS)。其思想是同时扰动 U_Y 、 U_X ，同时矫正 U_Y 、 U_X 中存在的噪声：

$$\begin{aligned}(U_X + \Delta U_X) \Psi &= U_Y + \Delta U_Y \implies \\ [-U_Y - \Delta U_Y, U_X + \Delta U_X] &\begin{bmatrix} 1 \\ \Psi \end{bmatrix} \\ &= ([-\Delta U_Y, \Delta U_X] + [-U_Y, U_X]) \begin{bmatrix} 1 \\ \Psi \end{bmatrix} \\ &= (\Delta U + U) z = 0\end{aligned}$$

于是 TLS 的解等价于

$$\begin{aligned}\min & \|\Delta U\|^2 \\ \text{s.t.} & (\Delta U + U) z = 0\end{aligned}$$

上面的问题通过下面的最小化问题得到 Ψ 的最小二乘解。记 $U_{ST} = [U_X | U_Y]$ ，寻找一个酉矩阵 F 。上面的 TLS 问题转化为

$$\begin{aligned} \min & \|U_{ST}F\|_F^2 \\ \text{s.t.} & F^H F = I \end{aligned}$$

我们发现只要使得 F 与 U_X 和 U_Y 都正交即可，则 F 可以从 $U_{ST}^H U_{ST}$ 的特征分解中得到。因为

$$\begin{aligned} U_{ST}F &= [U_X | U_Y] \begin{bmatrix} F_X \\ F_Y \end{bmatrix} \\ &= A^T F_X + A \Phi^T F_Y = 0 \\ \implies \Psi &= -F_X F_Y^{-1} \\ U_{ST}^H U_{ST} &= E \Sigma E^H \\ &= \begin{bmatrix} E_S & E_N \end{bmatrix} \begin{bmatrix} \Lambda_S & \\ & \Lambda_N \end{bmatrix} \begin{bmatrix} E_S^H \\ E_N^H \end{bmatrix} \end{aligned}$$

其中 Σ 是从大到小排列的特征值构成的对角矩阵，由 $U_{ST} [w_1, w_2, \dots, w_n] = [w_1, w_2, \dots, w_n] \text{diag}(\lambda)$ ， F 应该取后面的噪声子空间 E_N ，对应的 $\Lambda_N = \mathbf{0}$ ，可以达到最小化的目的。于是有

$$F = \begin{bmatrix} F_X \\ F_Y \end{bmatrix} = \begin{bmatrix} E_{12} \\ E_{22} \end{bmatrix} = E_N$$

其中，我们对矩阵 E 做了划分：

$$E = \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix}$$

因此有 $\Psi_{TLS} = -F_X F_Y^{-1} = -E_{12} E_{22}^{-1}$ ，然后求特征值，得到对应的到达角。

ESPRIT 算法仿真

```
% TLS_ESPRIT ALOGRITHM
% DOA ESTIMATION BY TLS_ESPRIT ALOGRITHM
clear variables;
close all; clc;

%信噪比(dB)
snr=-10:2:10;
lsnr = length(snr);
RMSE=zeros(1,lsnr);
test_num = 300;

source_number=1; % 信源数
sensor_number=8; % 原阵元数
m=7; % 子阵元数
N_x=1024; % 信号长度
snapshot_number=N_x; % 快拍数
w=pi/4; % 信号频率
l=2*pi*3e8/w; % 信号波长
d=0.5*l; % 阵元间距
array_distance=d; % 两组阵元轴线方向的间距

source_a= 50; % 两个信号的入射角度
source_doa= source_a *pi/180; % 两个信号的入射角度
```

```

% 阵列流型
A=[exp(-1i*(0:sensor_number-1)*d*2*pi*sin(source_doa)/l)].';

for isnr = 1:lsnr
    RMSE_i = zeros(1, test_num);
    for iter = 1:test_num
        % 仿真信号
        s=sqrt(10.^(snr(isnr)/10))*exp(1i*w*[0:N_x-1]);
        %x=awgn(s,snr);
        % 加了高斯白噪声后的阵列接收信号
        x=A*s+(1/sqrt(2))*(randn(sensor_number,N_x)+1i*randn(sensor_number,N_x));
        x1=x(1:m,:); % 子阵1接受的数据矢量
        x2=x(2:(m+1),:); % 子阵2接受的数据矢量

        % 对两个子阵的模型进行合并
        X = [x1; x2];
        R=X*X'/snapshot_number;
        % 对R进行奇异值分解
        [U,S,V]=svd(R);
        R=R-S(2*m,2*m)*eye(2*m);
        [U,S,V]=svd(R);
        Us=U(:,1:source_number);
        % disp(Us);
        Us1=Us(1:m,:);
        Us2=Us((m+1):2*m,:);
        %Us12=[Us1 Us2];
        % 形成矩阵Us12
        Us12=[Us1,Us2];
        % 对“Us12'*Us12”进行特征分解，得到矩阵E
        [E,Sa,Va]=svd(Us12'*Us12);
        % disp('E');
        % disp(E);
        % disp(Sa);
        % 将 E_4x4 分解为四个小矩阵
        E11=E(1,1);
        E12=E(1,2);
        E21=E(2,1);
        E22=E(2,2);
        % 按照公式得到旋转不变矩阵M
        M=-(E12*(inv(E22)));
        % disp('M');
        % disp(M);
        % 对得到的旋转不变矩阵进行特征分解
        [Vm,Dm]=eig(M);
        % disp(Dm);
        Dm=(diag(Dm)).';
        doa=-asin(angle(Dm)/pi)*180/pi;
        disp(doa);
        RMSE_i(iter)=sqrt((doa-source_a)^2);
    end
    RMSE(isnr) = sum(RMSE_i)/test_num;
end

%% plot
figure;
plot(snr,RMSE,'m-s');
legend('elevation','azimuth');

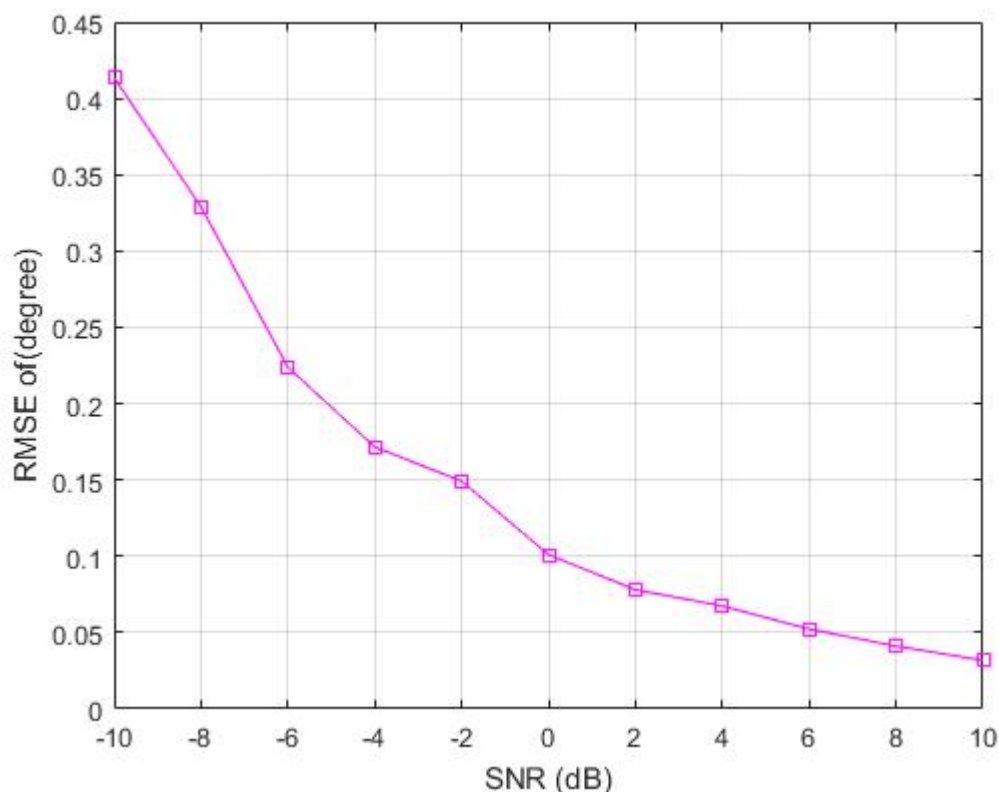
```

```

xlabel('SNR (dB)');
ylabel('RMSE of(degree)');
%axis([-100 100 -40 60]);
%set(gca, 'XTick',[-100:20:100]);
grid on;

```

运行效果如图：



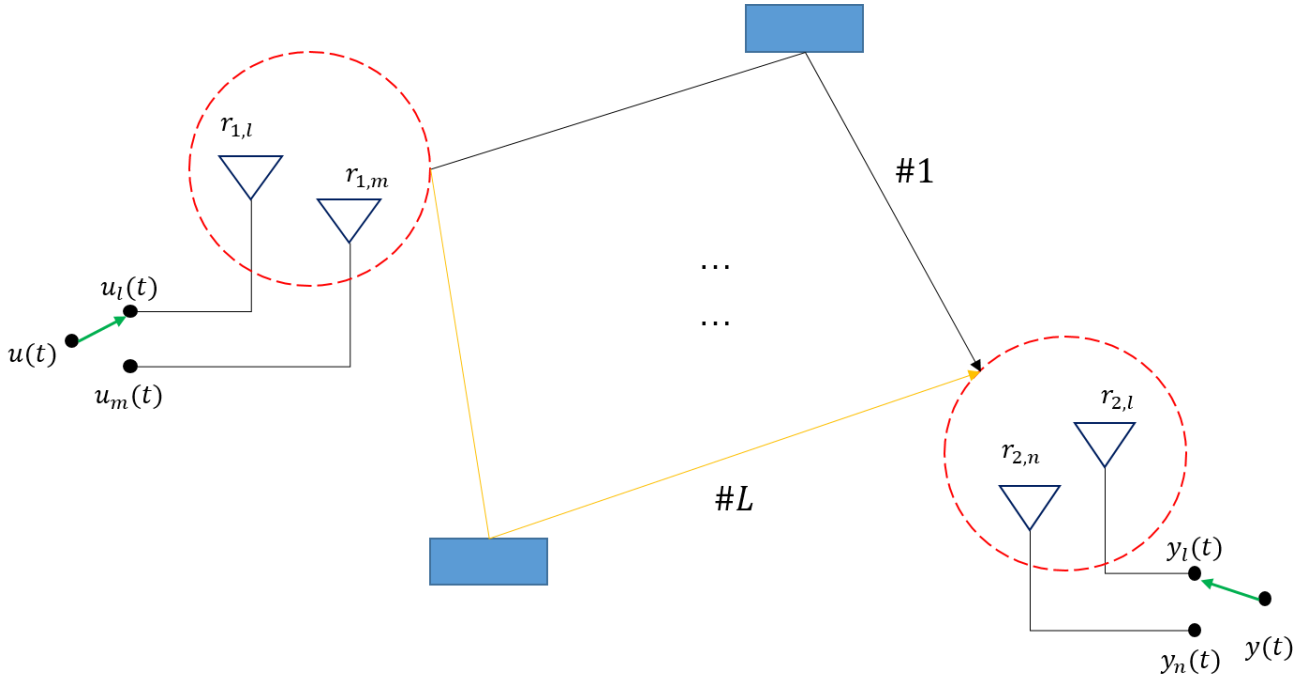
3. 期望最大算法

要想同时得到信道的多维参数信息，这里引入了 EM 算法来解决这一问题。EM 算法可以对无线传输信道中的无线电波进行**幅度、时延、离开角、到达角、多普勒谱**的联合多重参数估计。

系统模型

考虑如下的传播环境模型。发射端有 M 根天线组成的阵列 1，经过 L 条路径传输到达接收端，接收端有 N 根天线组成的阵列 2。在每条路径上，信号受到来自中途障碍物的干扰，发生不同程度的折射、反射、绕射等，使到达信号产生**多径现象**。

其中发射天线阵列由 M 个相对任意参考点 O_1 分别位于 $r_1, \dots, r_M \in R^3$ 的阵元组成，接收天线阵列由 N 个相对任意参考点 O_2 分别位于 $r_1, \dots, r_N \in R^3$ 的阵元组成。发射和接收天线阵列的阵元分别被限定在区域 R_1 和 R_2 中，这样我们假设的传播环境就满足了小尺度特性，对于我们后面的计算提供了便利。



该模型假设：

- **远场条件：**发射天线阵列和接收天线阵列的距离要求足够远，这样才能使得接收信号都可以看成**平面波**。同时保证发射天线轻微的移动几十个波长时，**每条径所经历的几何衰落路径基本保持不变**。
- **小尺度特性：**发射天线阵和接收天线阵都**限定在一定区域内**，使得发射端和接收端天线阵中的阵元的坐标都可以用三维矢量表示。信道的衰落特性符合**小尺度衰落特性**，即只考虑由于幅度、时延、离开角、到达角、多普勒频移等小尺度特性造成的信号衰落。

发射信号为 $\sqrt{\frac{E}{T}}\mathbf{u}(t)$ ，其中 T 和 E 分别表示周期和在周期内的能量， $\mathbf{u}(t) = [u_1(t), \dots, u_M(t)]^T$ 。

上述传播模型中假设有 L 条波到达接收机，接收机总共有 N 个天线阵元接收来波并对接收后的信号进行处理，这样第 n 个天线阵元所得的信道冲激响应可以表示为：

$$h_n(t) = \sum_{l=1}^L h_l e^{j2\pi v_l t} \mathbf{c}_2(\Omega_{2,l}) \mathbf{c}_1(\Omega_{1,l})^T \delta(t - \tau_l)$$

天线阵接收到的信号可以表示成：

$$\begin{aligned} Y_n(t) &= \sqrt{\frac{E}{T}} \mathbf{u}(t) * h_n(t) + N_n(t) \\ &= \sum_{l=1}^L \sqrt{\frac{E}{T}} h_l e^{j2\pi v_l t} \mathbf{c}_2(\Omega_{2,l}) \mathbf{c}_1(\Omega_{1,l})^T \mathbf{u}(t - \tau_l) + N_n(t) \\ &= \sum_{l=1}^L a_l e^{j2\pi v_l t} \mathbf{c}_2(\Omega_{2,l}) \mathbf{c}_1(\Omega_{1,l})^T \mathbf{u}(t - \tau_l) + N_n(t) \\ &= \sum_{l=1}^L \mathbf{s}(t; \boldsymbol{\theta}_l) + N_n(t) \end{aligned}$$

其中，经过第 l 条路径传输来的信号矩阵定义为：

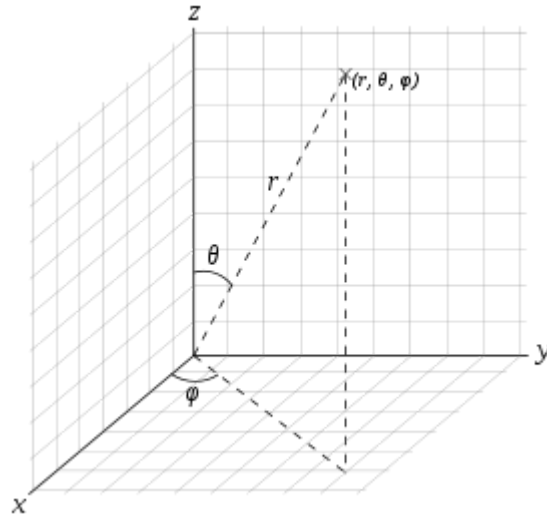
$$\begin{aligned} \mathbf{s}(t; \boldsymbol{\theta}_l) &= [s_1(t; \boldsymbol{\theta}_l), \dots, s_N(t; \boldsymbol{\theta}_l)]^T \\ &= \sum_{n=1}^N a_l e^{j2\pi v_l t} \mathbf{c}_2(\Omega_{2,l}) \mathbf{c}_1(\Omega_{1,l})^T \mathbf{u}(t - \tau_l) \end{aligned}$$

其中, $a_l = \sqrt{\frac{E}{T}} h_l$, $N_n(t) = N_{n\Re}(t) + jN_{n\Im}(t)$, $N_{n\Re}(t)$ 和 $N_{n\Im}(t)$ 均为具有单位频谱高度的零均值高斯白噪声过程, 且相互独立, 其双边功率谱密度为 $\frac{N_0}{2}$, 满足 $\mathbb{E}[N_n(t)^H N_n(t)] = 2\delta(t' - t)$ 。所以阵列 2 的接收信号为

$$\begin{aligned}\mathbf{Y}(t) &= [Y_1(t), \dots, Y_N(t)]^T \\ &= \sum_{l=1}^L \mathbf{s}(t; \boldsymbol{\theta}_l) + \sqrt{\frac{N_0}{2}} \mathbf{N}(t)\end{aligned}$$

其中, $\mathbf{N}(t) = [N_1(t), \dots, N_N(t)]^T$ 是 N 维复高斯白噪声向量值。 $\boldsymbol{\theta}_l = [\Omega_{1,l}, \Omega_{2,l}, \tau_l, v_l, a_l]$ 表示第 l 条路径的平面波参数矩阵。第 l 条路径的平面波由离开方向角 $\Omega_{1,l}$ 、到达方向角 $\Omega_{2,l}$ 、传播时延 τ_l 、多普勒频移 v_l 和复振幅 a_l 来表征。因此, 信道参数提取算法的核心就在于准确有效地估计这 5 个信道参数的信息。

为了简洁表达, 我们定义 $\mathbf{s}(t; \boldsymbol{\theta}) = \sum_{l=1}^L \mathbf{s}(t; \boldsymbol{\theta}_l)$, 其中 $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_L]^T$ 。为了方便描述到达角和离开角的方向, 在这里引入**球坐标**。



首先初始点设定为 O 处, 以这个点为球心画出一个半径为 1 的球, 确定一个方向, 作为单位向量 Ω 。球面坐标系的**单位向量**为:

$$\Omega = [\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta]^T$$

其中角度 $\phi \in [-\pi, \pi)$ 为单位矢量 Ω 的水平方向角, 角度 $\theta \in [0, \pi]$ 表示单位矢量 Ω 的垂直方向角。于是有

$$\begin{aligned}\mathbf{c}_2(\Omega_2) &= [f_{2,1}(\Omega_2) \exp\{j2\pi\lambda_0^{-1}(\Omega_2 \cdot \mathbf{r}_{2,1})\}, \dots, f_{2,N}(\Omega_2) \exp\{j2\pi\lambda_0^{-1}(\Omega_2 \cdot \mathbf{r}_{2,N})\}]^T \\ \mathbf{c}_1(\Omega_1) &= [f_{1,1}(\Omega_1) \exp\{j2\pi\lambda_0^{-1}(\Omega_1 \cdot \mathbf{r}_{1,1})\}, \dots, f_{1,M}(\Omega_1) \exp\{j2\pi\lambda_0^{-1}(\Omega_1 \cdot \mathbf{r}_{1,M})\}]^T\end{aligned}$$

上述表达式中 λ_0 为传播波的波长, $f_{1,m}(\Omega_1)$ 为发射天线阵列第 m 个阵元的复电场模式, $f_{2,n}(\Omega_2)$ 为接收天线阵列第 n 个阵元的复电场模式, (\cdot) 为内积符号。

由观测时间 PT 内的观测值 $\mathbf{Y}(t) = y(t)$ 可得 $\boldsymbol{\theta}$ 的对数似然函数:

$$L(\boldsymbol{\theta}; y) = \frac{1}{N_0} \left[\underbrace{2 \int_{PT} \Re \{ \mathbf{s}^H(t; \boldsymbol{\theta}) y(t) \} dt}_{G_1} - \underbrace{\int_{PT} \|\mathbf{s}(t; \boldsymbol{\theta})\|^2 dt}_{G_2} \right]$$

其中, $\Re(\cdot)$ 表示取实部。 $\boldsymbol{\theta}$ 的最大似然估计的结果就是使对数似然函数 $L(\boldsymbol{\theta}; y)$ 取最大值的向量, 即

$$\hat{\boldsymbol{\theta}}_{ML}(y) = \arg \max_{\boldsymbol{\theta}_l} \{L(\boldsymbol{\theta}; y)\}$$

对于信道中多径数目较大时, 上式是关于 $L(\boldsymbol{\theta}; y)$ 的非线性函数, 计算机无法直接对其进行求解。

EM 算法原理

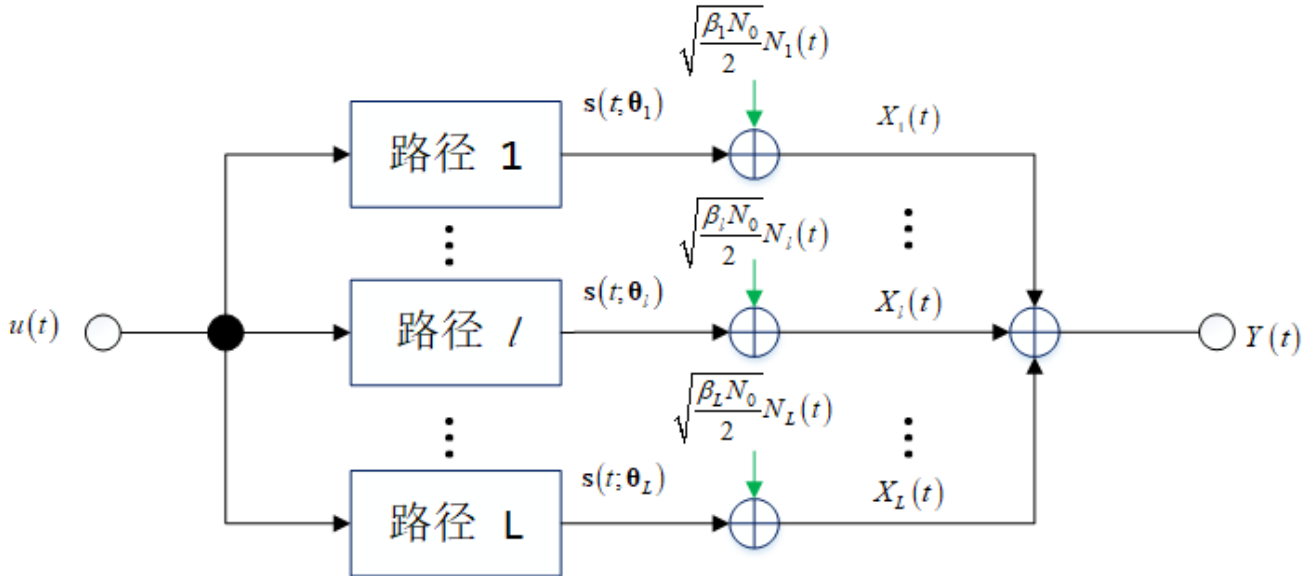
EM 算法的推导基于两个主要概念：完全数据（不可观测）和不完全数据（可观测）。其思想是：你知道一些东西（观察到的数据），你不知道一些东西（观察不到的数据），如果你想知道那些不了解的东西，采取的方法就是根据一些假设 (parameter) 先猜 (E-step)，把不知道的事都猜出来，基于知道的和猜测的数据，假装你全知道了，然后根据猜测的数据，你反思 (M-step) 一下，更新假设，使得观察到的数据更加可能。然后再猜，再反思，最后就得到可以解释整个数据的假设了。

EM 算法的核心思想是将需要估计的参数集 θ 分割成参数子集 $\theta_1, \dots, \theta_L$ 顺序估计。

定义 $Y(t)$ 为实际数据的**不完整可观测值**，可表示成一组**完整但不可观测**的数据的函数 $[X_1(t), \dots, X_L(t)]$ ，其中 $X_l(t) = s(t; \theta_l) + \sqrt{\beta_l} \sqrt{\frac{N_0}{2}} N_l(t)$ ，在这里**重新定义** $N(t) = [\sqrt{\beta_1} N_1(t), \dots, \sqrt{\beta_L} N_L(t)]$ 是方差为 β_l 的高斯平稳噪声，且 $N_l(t)$ 之间相互独立。其中 β_l 为正值，且 $\sum_{l=1}^L \beta_l = 1$ 。则不完整数据可以表示为完整数据的函数：

$$Y(t) = \sum_{l=1}^L X_l(t) = \sum_{l=1}^L s(t; \theta_l) + \sum_{l=1}^L \sqrt{\beta_l} \sqrt{\frac{N_0}{2}} N_l(t)$$

可测数据与不可测数据转换示意图：



我们**假设完全数据** X **是可以观测的**。由于 $X_1(t), \dots, X_L(t)$ 相互独立，故对第 l 条路径的传播参数 θ_l 的估计与其它路径无关。由观测时间 PT 内的观测值 $X_l(t) = x_l(t)$ ，可得 θ_l 的对数似然函数：

$$L(\theta_l; x_l) = \frac{1}{\beta_l N_0} \left[\underbrace{2 \int_{PT} \Re \{s^H(t; \theta_l) x_l(t)\} dt}_{G_1} - \underbrace{\int_{PT} \|s(t; \theta_l)\|^2 dt}_{G_2} \right]$$

对于已知观测值 $X_l(t) = x_l(t)$ ， θ_l 的最大似然估计为：

$$(\hat{\theta}_l)_{ML}(x_l) = \arg \max_{\theta_l} \{L(\theta_l; x_l)\}$$

因为 $X_l(t)$ 是不可观测的完全数据，这里采用可以观测的不完全数据 $Y(t) = y(t)$ 以及 θ 的上一次估计值 $\hat{\theta}'$ 对其进行估计。于是可以利用已知的 $Y(t) = y(t)$ 以及假设的 $\theta = \hat{\theta}'$ 从而计算 $X_l(t)$ 条件期望：

$$x_l(t; \hat{\theta}') = \mathbb{E}_{\hat{\theta}'} [X_l(t) | Y(t) = y(t)], \quad l = 1, 2, \dots, L$$

此处， $\mathbb{E}_{\hat{\theta}'}(\cdot)$ 指在假设参数值 θ 情况下的期望。利用条件期望所得的 $x_l(t; \hat{\theta}')$ 作为 $X_l(t)$ 的观测值，然后通过最大似然估计对传播波参数矢量 θ_l 进行再估计：

$$\hat{\theta}_l'' = (\hat{\theta}_l)_{ML}(\hat{x}_l(t; \hat{\theta}')), \quad l = 1, 2, \dots, L$$

总结以上两个公式就对应着 EM 算法的操作：

EM 算法	公式
E (求期望)	$x_l(t; \hat{\theta}') = \mathbb{E}_{\hat{\theta}'}[X_l(t) Y(t) = y(t)], \quad l = 1, 2, \dots, L$
M (最大化)	$\hat{\theta}_l'' = (\hat{\theta}_l)_{ML}(\hat{x}_l(t; \hat{\theta}')), \quad l = 1, 2, \dots, L$

对于任意初始值 $\hat{\theta}(0)$ ，多次交替迭代执行 E 和 M 步骤。对于第 μ 次迭代，由初始值 $\hat{\theta}(\mu - 1)$ 去求得每一径传播波的参数 $\hat{\theta}_l(\mu)$ 。迭代到收敛为止。

算法的 M 步骤，根据函数 $x_l(t; \hat{\theta}')$ 的特殊结构，其参数向量 $\theta_l = [\Omega_{1,l}, \Omega_{2,l}, \tau_l, v_l, a_l] = [\bar{\theta}_l, a_l]$ 可以简化为以下几个步骤。首先将 a_l 固定，求函数另外 4L 个维度的联合最大化，为了进一步简化该步骤，可以将联合最大化变成 4 个参数分别的最大化。求得 $\hat{\Omega}_{1,l}'', \hat{\Omega}_{2,l}'', \hat{\tau}_l'', \hat{v}_l''$ 。之后可以通过以 $\hat{\Omega}_{1,l}'', \hat{\Omega}_{2,l}'', \hat{\tau}_l'', \hat{v}_l''$ 为函数的闭合表达式得到 \hat{a}_l'' 。将复振幅 a_l 的值代入到 $L(\theta_l; x_l)$ 中，调整 M 步骤的公式如下：

$$\begin{aligned} (\hat{\bar{\theta}}_l)_{ML}(x_l) &= \arg \max_{\bar{\theta}_l} \{z(\bar{\theta}_l; x_l)\}, \\ (\hat{a}_l)_{ML}(x_l) &= \frac{z(\bar{\theta}_l; x_l) \Big|_{\bar{\theta}_l = (\hat{\bar{\theta}}_l)_{ML}(x_l)}}{\left| \mathbf{c}_2 \left((\hat{\Omega}_{2,l})_{ML} \right) \right| \cdot \left| \mathbf{c}_1 \left((\hat{\Omega}_{1,l})_{ML} \right) \right| I \cdot P \cdot T_{SC}}. \end{aligned}$$

其中相关文献中介绍的无线信道探测仪在接收端和发射端都配备了射频天线开关，这里采用时分多路复用(TDM)模式进行无线信道的探测。用 I 表示发射端循环周期的个数，将 I 个循环周期定义为一个快拍。发射端发送周期为 T_s 且功率为 P 的探测信号。对于接收端，由于射频天线开关的作用，接收天线阵列的 N 个阵元也相继对接收信号进行扫描，每个阵元的扫描时间为 T_{SC} 。

$$\begin{aligned} z(\bar{\theta}_l; x_l) &= \mathbf{f}(\bar{\theta}_l)^H \mathbf{D}(\Omega_{1,l}, \Omega_{2,l}) \mathbf{f}(\bar{\theta}_l) \\ &= \tilde{\mathbf{c}}_2(\Omega_{2,l})^H \mathbf{X}_l(\tau_l, v_l) \tilde{\mathbf{c}}_1(\Omega_{1,l})^* \end{aligned}$$

其中，

$$\begin{aligned} \mathbf{f}(\bar{\theta}_l) &= \begin{bmatrix} \mathbf{c}_{2,1}(\Omega_{2,l})^H \mathbf{X}_l(\tau_l, v_l) \mathbf{c}_{1,1}(\Omega_{1,l})^* \\ \mathbf{c}_{2,1}(\Omega_{2,l})^H \mathbf{X}_l(\tau_l, v_l) \mathbf{c}_{1,2}(\Omega_{1,l})^* \\ \mathbf{c}_{2,2}(\Omega_{2,l})^H \mathbf{X}_l(\tau_l, v_l) \mathbf{c}_{1,1}(\Omega_{1,l})^* \\ \mathbf{c}_{2,2}(\Omega_{2,l})^H \mathbf{X}_l(\tau_l, v_l) \mathbf{c}_{1,2}(\Omega_{1,l})^* \end{bmatrix}, \\ \tilde{\mathbf{c}}_2(\Omega_{2,l}) &= \frac{\mathbf{c}_2(\Omega_{2,l})}{|\mathbf{c}_2(\Omega_{2,l})|}, \\ \tilde{\mathbf{c}}_1(\Omega_{1,l}) &= \frac{\mathbf{c}_1(\Omega_{1,l})}{|\mathbf{c}_1(\Omega_{1,l})|}. \end{aligned}$$

$\mathbf{X}_l(\tau_l, v_l)$ 是 $N \times M$ 维的矩阵，其中每一个元素如下所示：

$$x_{l,n,m}(t; \tau_l, v_l) = \sum_{i=1}^I \exp\{-j2\pi v_l t_{i,n,m}\} \cdot \int_0^{T_{SC}} u(t - \tau_l)^* \exp\{-j2\pi v_l t\} x_l(t + t_{i,n,m}) dt$$

其中， $\mathbf{D}(\Omega_{1,l}, \Omega_{2,l}) = [\mathbf{c}_2(\Omega_{2,l})^H \mathbf{c}_2(\Omega_{2,l})] \otimes [\mathbf{c}_1(\Omega_{1,l})^H \mathbf{c}_1(\Omega_{1,l})]$ 。为了保证算法的 M 步骤中的对于参数矩阵的估计有解，需要满足矩阵 $\mathbf{D}(\Omega_{1,l}, \Omega_{2,l})$ 的非奇异值条件，即行列式不为 0：

$$\det(\mathbf{D}(\Omega_{1,l}, \Omega_{2,l})) = \prod_{k=1}^2 \left(\left| \mathbf{c}_{k,1}(\Omega_k) \right|^2 \left| \mathbf{c}_{k,2}(\Omega_k) \right|^2 - \left| \mathbf{c}_{k,2}(\Omega_k)^H \mathbf{c}_{k,1}(\Omega_k) \right|^2 \right)^2$$

发现上式 $\det(\mathbf{D}(\Omega_{1,l}, \Omega_{2,l})) \geq 0$, 当且仅当 $\mathbf{c}_{k,1}(\Omega_k)$ 和 $\mathbf{c}_{k,2}(\Omega_k)$ 线性相关时, 等号才成立。故矩阵 $\mathbf{D}(\Omega_{1,l}, \Omega_{2,l})$ 是非奇异矩阵的充要条件为: 对于任意的 Ω_k , ($k = 1, 2$), $\mathbf{c}_{k,1}(\Omega_k)$ 和 $\mathbf{c}_{k,2}(\Omega_k)$ 都是线性不相关的。

SAGE 算法

根据似然函数

$$L(\boldsymbol{\theta}_l; x_l) = \frac{1}{\beta_l N_0} \left[\underbrace{2 \int_{PT} \Re \{ \mathbf{s}^H(t; \boldsymbol{\theta}_l) x_l(t) \} dt}_{G_1} - \underbrace{\int_{PT} \|\mathbf{s}(t; \boldsymbol{\theta}_l)\|^2 dt}_{G_2} \right]$$

其中,

$$\begin{aligned} G_1 &= a_l^* \mathbf{c}_2(\Omega_{2,l})^H \mathbf{X}_l(\tau_l, v_l) \mathbf{c}_1(\Omega_{1,l})^*, \\ G_2 &= |a_l|^2 \left| \mathbf{c}_2(\Omega_{2,l}) \right|^2 \left| \mathbf{c}_1(\Omega_{1,l}) \right|^2 IPT_{SC}. \end{aligned}$$

于是, 代入整理可得似然函数

$$L(\boldsymbol{\theta}_l; x_l) \propto 2 \left| \mathbf{c}_2(\Omega_{2,l}) \right| \cdot \left| \mathbf{c}_1(\Omega_{1,l}) \right| \Re \{ a_l^* z(\bar{\boldsymbol{\theta}}_l; x_l) \} - |a_l|^2 \left| \mathbf{c}_2(\Omega_{2,l}) \right|^2 \left| \mathbf{c}_1(\Omega_{1,l}) \right|^2 IPT_{SC}.$$

SAGE 算法是针对 EM 算法**每次迭代都要更新所有参数**这一缺陷所提出的一种优化扩展算法。SAGE 算法的主要思想是: 在 EM 算法的基础上将参数矩阵空间集分割成维数较小的参数子集, 每次迭代只更新部分参数子集, 而其余补集保持不变。具体的做法为: 将每一径的参数矩阵分割为 6 个参数子集:

$$\begin{aligned} \boldsymbol{\theta}_l &= [\Omega_{1,l}, \Omega_{2,l}, \tau_l, v_l, a_l] \\ &\implies \\ &\quad \{\tau_l, a_l\}, \\ &\quad \{\theta_{2,l}, a_l\}, \\ &\quad \{\phi_{2,l}, a_l\}, \\ &\quad \{\theta_{1,l}, a_l\}, \\ &\quad \{\phi_{1,l}, a_l\}, \\ &\quad \{v_l, a_l\}. \end{aligned}$$

对于不可观测的完全数据 $X_l(t)$ 的估计与 $\hat{\boldsymbol{\theta}}'$ 是无关的。所以对参数空间集进行分割不影响求期望步骤的进行, 每个子集的求期望步骤是完全一致的。在迭代的过程中, 每次只更新其中一个子集, 其余补集都保持初始值不变, 依次更新所有子集后一条路径一次迭代完成。对所有路径完成一次迭代即完成了一个迭代周期。只更新子集的做法可以减小 Fisher 信息, 算法的运算复杂度得到显著降低, 收敛速度明显加快。一次迭代所得到的参数矩阵的值如下所示:

$$\begin{aligned} \hat{\tau}_l'' &= \arg \max_{\tau_l} z(\hat{\phi}_{1,l}', \hat{\theta}_{1,l}', \hat{\phi}_{2,l}', \hat{\theta}_{2,l}', \tau_l, \hat{v}_l', \hat{x}_l), \\ \hat{\theta}_{2,l}'' &= \arg \max_{\theta_{2,l}} z(\hat{\phi}_{1,l}', \hat{\theta}_{1,l}', \hat{\phi}_{2,l}', \theta_{2,l}, \hat{\tau}_l'', \hat{v}_l', \hat{x}_l), \\ \hat{\phi}_{2,l}'' &= \arg \max_{\phi_{2,l}} z(\hat{\phi}_{1,l}', \hat{\theta}_{1,l}', \phi_{2,l}, \hat{\theta}_{2,l}'', \hat{\tau}_l'', \hat{v}_l', \hat{x}_l), \\ \hat{\theta}_{1,l}'' &= \arg \max_{\theta_{1,l}} z(\hat{\phi}_{1,l}', \theta_{1,l}, \hat{\phi}_{2,l}'', \hat{\theta}_{2,l}'', \hat{\tau}_l'', \hat{v}_l', \hat{x}_l), \\ \hat{\phi}_{1,l}'' &= \arg \max_{\phi_{1,l}} z(\phi_{1,l}, \hat{\theta}_{1,l}'', \hat{\phi}_{2,l}'', \hat{\theta}_{2,l}'', \hat{\tau}_l'', \hat{v}_l', \hat{x}_l), \\ \hat{v}_l'' &= \arg \max_{v_l} z(\hat{\phi}_{1,l}'', \hat{\theta}_{1,l}'', \hat{\phi}_{2,l}'', \hat{\theta}_{2,l}'', \hat{\tau}_l'', v_l, \hat{x}_l). \end{aligned}$$

最后 $\hat{a}_l'' = \left[\left| \mathbf{c}_2 \left((\hat{\phi}_{2,l}'', \hat{\theta}_{2,l}'') \right) \right| \cdot \left| \mathbf{c}_1 \left((\hat{\phi}_{1,l}'', \hat{\theta}_{1,l}'') \right) \right| IPT_{SC} \right]^{-1} \cdot z(\hat{\phi}_{1,l}'', \hat{\theta}_{1,l}'', \hat{\phi}_{2,l}'', \hat{\theta}_{2,l}'', \hat{\tau}_l'', \hat{v}_l'', \hat{x}_l)$, SAGE 算法具体步骤:

■ E 步骤

因为 $X_l(t)$ 不可观测, 所以对数似然函数也不可计算。那么关键步骤是基于 $Y(t) = y(t)$ 估计 $X_l(t)$ 的条件期望:

$$\begin{aligned} x_l(t; \hat{\theta}') &= \mathbb{E}_{\hat{\theta}'} \left[X_l(t) | Y(t) = y(t) \right] \\ &= y(t) - \sum_{l'=1, l' \neq l}^L s(t; \hat{\theta}_{l'}') \end{aligned}$$

■ M 步骤

$$\begin{aligned} (\hat{\theta}_l)_{ML}(x_l) &= \arg \max_{\bar{\theta}_l} \{ z(\bar{\theta}_l; x_l) \}, \\ (\hat{a}_l)_{ML}(x_l) &= \frac{z(\bar{\theta}_l; x_l) \Big|_{\bar{\theta}_l = (\hat{\theta}_l)_{ML}(x_l)}}{\left| \mathbf{c}_2 \left((\hat{\Omega}_{2,l})_{ML} \right) \right| \cdot \left| \mathbf{c}_1 \left((\hat{\Omega}_{1,l})_{ML} \right) \right| I \cdot P \cdot T_{SC}}. \end{aligned}$$

从上可以看出, SAGE 算法和 EM 算法的 E 步骤和 M 步骤的公式在形式上是完全一样的。SAGE 算法和 EM 算法的求期望过程是完全一致的, 但是 SAGE 算法在 EM 算法基础上, 对其 M 步骤进行了改进: 对于参数矩阵的估计是将其参数空间进行分割, 将多维联合参数估计转化成一维参数估计。

SAGE 算法仿真