

# 统计推断在 ADC 系统中的应用

组号：60 组员：宋凯敏 F1203021 5120309575

摘要：

本文是上海交通大学电子信息与电气工程学院课程设计《统计推断在 ADC 系统中的应用》的课程论文。本报告以 ftp 的实验结果为研究数据，探讨统计推断在数模模数转换中的应用。我们运用 Matlab 数学软件并使用统计推断的方法，通过 469 组数据对系统的输入输出关系进行研究。

关键词：统计推断、Matlab、特征点、曲线拟合、模拟退火算法。

## Application of Statistical Inference in ADC Inverting System

**ABSTRACT:** This article is for Course Design-Application of Statistical Inference in ADC System, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University. This report takes the results in ftp as research data, to explore the application of statistical inference in the digital to analog conversion. We use statistical inference, in combination with Matlab, to research into the relationship between the input and output, based on the given data.

**Key words:** statistical inference、Matlab、Feature points、curve fitting、simulated annealing algorithm

# 1 引言

本课题研究的内容为输入与输出之间的函数关系，由于两者并不存在确定的函数关系，因而不仅需要通过研究大量的数据（469 组实验数据）还要寻找适合的拟合函数、筛选相关数据，运算量较大，并且较为难以实现。因此，本课题旨在找出我们能够找出的误差较小的拟合函数，是的所有样本的共同点在这一函数中得以体现出来。

## 2 课程问题

基本问题的最终目标是找到 6 个特征点，是根据这 6 个点拟合的数据按照提供的函数计算后，所有的样本的越低越好。由题条件，成本的结果按照以下规则：

( $S_i$  为第  $i$  组的得分， $\hat{Y}_{i,j}$  表示第  $i$  组数据的第  $j$  点测得的输出， $Y_{i,j}$  表示根据拟合函数求得的输出)

$$s_{i,j} = \begin{cases} 0 & \text{if } |\hat{y}_{i,j} - y_{i,j}| \leq 1 \\ 1 & \text{if } 1 < |\hat{y}_{i,j} - y_{i,j}| \leq 2 \\ 2 & \text{if } 2 < |\hat{y}_{i,j} - y_{i,j}| \leq 3 \\ 4 & \text{if } 3 < |\hat{y}_{i,j} - y_{i,j}| \leq 5 \\ 10 & \text{if } |\hat{y}_{i,j} - y_{i,j}| > 5 \end{cases}$$

$$S_i = \sum_{j=1}^{51} s_{i,j} + 20N_i$$

最终的成本为

$$C = \frac{1}{M} \sum_{i=1}^M S_i$$

最终的总成本为

而第一个点与第六个点的选取，我们固定为第 1 点与第 51 点。

### 3 关于函数拟合

输入和输出之间虽然没有很明显的线性关系，但是其依旧具有较为平滑的曲线关系。

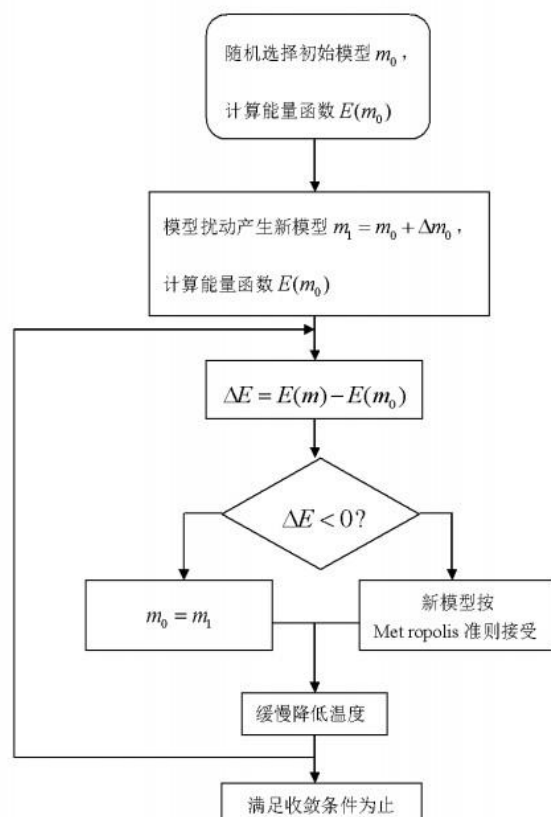
因此，我们首先尝试采用启发式 heuristic 搜索算法中的模拟退火算法对 6 个事前观测点进行选取以及分段多项式拟合来对各段函数进行拟合。

#### 3.1 模拟退火算法

##### 3.1.1 模拟退火算法简介

模拟退火算法来源于固体退火原理，将固体加温至充分高，再让其徐徐冷却，加温时，固体内部粒子随温升变为无序状，内能增大，而徐徐冷却时粒子渐趋有序，在每个温度都达到平衡态，最后在常温时达到基态，内能减为最小。

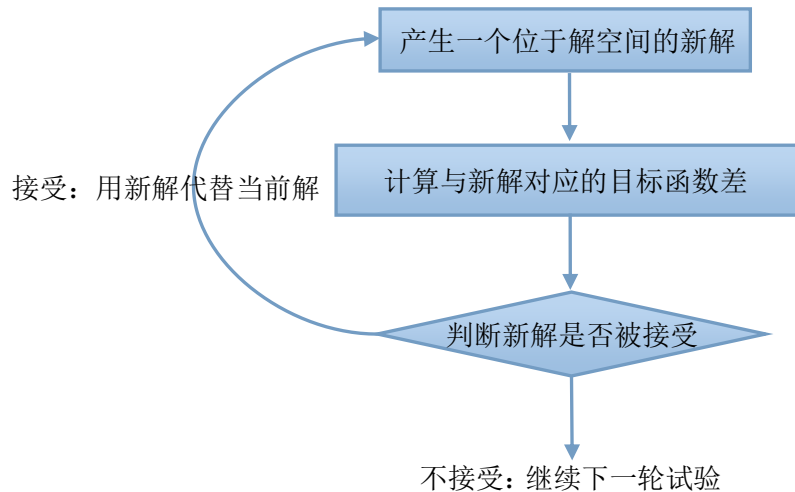
模拟退火算法：由初始解  $i$  和控制参数初值  $t$  开始，对当前解重复“产生新解→计算目标函数差→接受或舍弃”的迭代，并逐步衰减  $t$  值，算法终止时的当前解即为所得近似最优解，这是基于蒙特卡罗迭代求解法的一种启发式随机搜索过程。退火过程由冷却进度表(Cooling Schedule)控制，包括控制参数的初值  $t$  及其衰减因子  $\Delta t$ 、每个  $t$  值时的迭代次数  $L$  和停止条件  $S$ 。



### 3.1.1 模拟退火算法在本报告中的应用

模拟退火算法可以分解为解空间、目标函数和初始解三部分。

我们根据模拟退火算法，在本次报告中新解的产生和接受可以用一下流程图表示：



关于以上流程图的解释：

- 1、生成函数的初解：初解的生成借用随机函数在 469 组的数据中产生。
- 2、判断新解是否被接受：应用 Metropolis 准则：若 $<0$  则接受为新的当前解，否则以概率  $\exp(-\Delta t/T)$  接受作为新的当前解。
- 3、新解生成函数：扰动的产生依旧借助于随机函数在原数据组中进行随机取样，每次改变 4 个点中的一个。

（对于这观测点的取值范围我定为： ）

序号	取值范围
1	1
2	7~27
3	13~33
4	19~39
5	25~45
6	51

- 4、终止条件：目标函数的成本函数结果小于等于 130（其中 120 为测定成本，10 为误差成本）

## 3.2 分段多项式拟合

由于整体系统特征与其电路系统中晶体管元件特性密切相关，晶体管工作特性可分为前后衔接的连续区间：截止区、线性区和饱和区，在跨越区间是物理特性是连续而非跳跃间断的；我们假定整体系统特性也会呈现前后衔接的  $N$  个区间，每个徐建适合分别用一段多项式曲线进行拟合，前后区间对应的多项式曲线在区间衔接处取值连续。

### 3.2.1 区间分段点的选取

我通过各个点处所对应的平均曲率变化来寻找区间分段点。

我通过对数据库中若干组解的取平均值后，可以得到近似的曲率的区间分为一个分段，这样我们就可以在间隔为 6 的数组中找到有限个分段区间，进而可以得到相对精确的分段拟合函数，如之前退火算法中介绍过的：

序号	取值范围
1	1
2	7~27
3	13~33
4	19~39
5	25~45
6	51

### 3.2.2 每个区间分段中函数的拟合

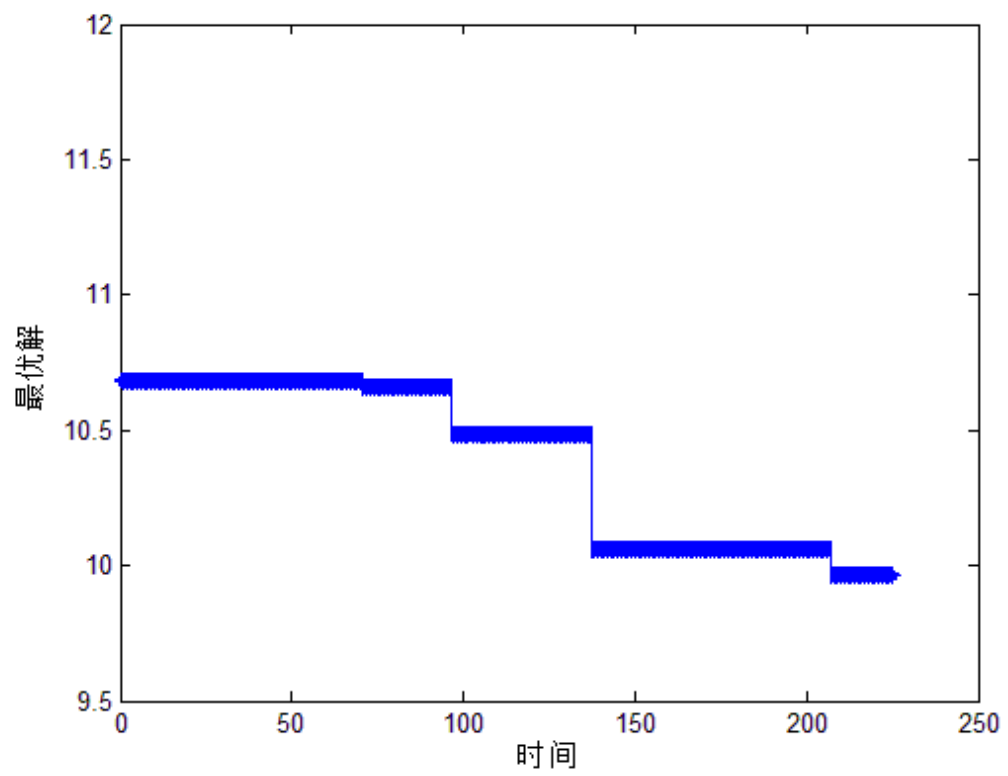
一般意义而言，最佳的拟合函数是个不可能达到的要求，部分样本自身数据的误差会导致拟合效果极差，基本不可能出现对于所有的数据点都极为符合的情况。往届同学的研究也反映了这一点，在此条件下，连续多次的拟合尝试都没有能达到最佳的效果。在后期的继续尝试中，我们尝试了基于最小二乘法的多项式拟合，在每一个分段当中，有的分段线性较好的可以撇除离散点对其拟合的影响，而对于曲线类型的分段，我们原本预期将进行三次以及四次多项式拟合的尝试。

最终发现，三次的拟合就能达到目标的效果，相比于二次的拟合有了极为明显的提升，相对算法也较为简单。因此我们对于中间的分段进行三次的拟合，应用与中间的分段，而对于两边的分段仍旧应用二次拟合。（如下所示）

分段	拟合方式
1——2	通过 1,2,3 点的二次拟合
2——3	通过 1,2,3,4 点的三次拟合
3——4	通过 2,3,4,5 点的三次拟合
4——5	通过 3,4,5,6 点的三次拟合
5——6	通过 4,5,6 点的二次拟合

最终的出最优点为：1、9、21、32、44、51（成本平均值为 9.8290）

最优解分布如图所示：



## 4 参考文献

- [1] 上海交大电子工程系, 统计推断在 ADC 系统中的应用课程讲义
- [2] Matlab 在数学建模中的应用 卓金武 北京航空航天大学出版社
- [3] 模拟退火算法机理研究 陈华根, 吴建生, 王家林, 等。同济大学学报: 自然科学版, 2004,32 (6): 802-805
- [4] 百度百科

## 6 附录（代码）

```
function main()
clc;
clear all;
close all;
global data
global data_size
tic;
data = dlmread('20141010dataform.csv')
data_size =size( data );
select2(data);
```

```
%-----
function select2(data)
si=[ 1      11      21      31      41      51 ];
snew=si;
smax=zeros(6);
[n,m]=size(data);
D=zeros(n/2,m);
U=D;
Ui=U;
score=zeros(n/2,1);
k=1:2:n;
D=data(k,:);
k=2:2:n;
U=data(k,:);
T=5000;
Tmin=0.05;
r=0.95;
fenold=targetfun(si);
disp(fenold);
scoremin=fenold;
times=1;

%开始模拟退火%
tic

while (T>Tmin)
    change=round((rand(1,6)-0.5)*2);
    %   if sum(abs(change))>4
    %       continue
    %   end
    snew=si+change;
```



```

if
    (sum(snew==0) || sum(snew==52) || abs(snew(1)-snew(2))<6 || abs(snew(2)-snew(3))<6 || ...)

    abs(snew(2)-snew(3))<6 || abs(snew(3)-snew(4))<6 || abs(snew(4)-snew(5))<6 || ...
        abs(snew(5)-snew(6))<6)
continue
end
disp(T);
disp(snew);
disp(scoremin);
fennew=targetfun(snew);
ds=fennew-fenold;
if ds>0
    si=snew;
    fenold=fennew;
    if fennew<scoremin
        smax=si;
        scoremin=fennew;
    end
else
    if exp(1000*ds/T)>rand
        si=snew;
        fenold=fennew;
    end
end
T=T*r;
a(times)=scoremin;
times=times+1;

end
times=times-1;
toc
disp(smax)
disp(fenold)
disp('模拟退火结束')

figure(1);
hand1=plot(1:times,a);
set(hand1,'linestyle','-','linewidth',1.8,'marker','*','markersize',6)
xlabel('时间');ylabel('最优解');
hold on;

```

```

function seln=selection(population,cumsump);
%从种群中选择两个个体
for i=1:2
    r=rand; %产生一个随机数
    prand=cumsump-r;
    j=1;
    while prand(j)<0
        j=j+1;
    end
    seln(i)=j; %选中个体的序号
end

%-----
global Data
global Data_size
global flag
Data = dlmread('20141010dataform.csv')
Data_size =size( Data );

individual=[1 6 20 36 45 51]
individual=sort(individual);
%解决相同元素问题
pc=1;
while pc>0
    pc=-1;
    individual=sort(individual);
    for i=2:6
        flag=round(individual(i)-individual(i-1));
        if flag==0
            individual(i)=round(individual(i)+1);
            if round(individual(i))==52
                individual(i)=1;
                pc=1;
                individual=sort(individual);

                break
            end
        else
            individual(i)=individual(i);
        end
    end
end
end

```

```

end

individual=sort(individual);

Score = zeros( 1, Data_size(1)/2 );
for i = 1 : Data_size(1)/2
    Y = interp1( Data( 2*i-1 , individual(1, : ) ),Data( 2*i,
individual(1, : ) ),Data( 2*i-1 ,2:51 ),'spline');
    dert_Y = abs( Y(:, :) - Data( 2*i, 2:51 ) );
    for j = 1:( Data_size(2) - 2 )
        if dert_Y( j ) <= 1
            Score( 1, i ) = Score( 1, i ) ;
        elseif dert_Y( j ) <= 2
            Score( 1, i ) = Score( 1, i ) + 1;
        elseif dert_Y( j ) <= 3
            Score( 1, i ) = Score( 1, i ) + 2;
        elseif dert_Y( j ) <= 5
            Score( 1, i ) = Score( 1, i ) + 4;
        else
            Score( 1, i ) = Score( 1, i ) + 10;
        end
    end
end
end
y=mean( Score(1, :) );

%-----
function [Fitvalue,cumsump]=fitnessfun(population);
global BitLength
global boundsbegin
global boundsend

popsize=size(population,1); %有 popsize 个个体

for b=1:popsize
    x=transform2to10(population(b, :));
    xx=round(boundsbegin+x*(boundsend-boundsbegin)/(power(2 ,7)-1));%
    二进制转十进制, 然后再转化为[1,51]区间的实数

    Fitvalue(b)=targetfun(xx);
end

%给适应度函数加上一个大小合理的数以便保证种群适应值为正数
Fitvalue=Fitvalue'+230;

```

```

%计算选择概率
fsum=sum(Fitvalue);
Pperpopulation=Fitvalue/fsum;
%计算累积概率
cumsum(1)=Pperpopulation(1);
for i=2:popsize
    cumsum(i)=cumsum(i-1)+Pperpopulation(i);
end

cumsum=cumsum';

%-----
function c=fitting_try(x,y,m)

%x 为数据点的横坐标
%y 为数据点的纵坐标
%m 为多项式最高次数
%返回结果为 F，拟合多项式函数

n=length(x);
b=zeros(1,m+1);
f=zeros(n,m+1);%f 为正规方程的系数,初始为 0_try
for k=1:m+1
    f(:,k)=x'.^(k-1);
end
a=f'*f;
b=f'*y';
%解方程,得到多项式由高到低的系数所构成的向量 c
c=a\b;
c=flipud(c);%翻转
%以上用了最小二乘法
end

%-----
function [population]=makesureDifferent(population);
popsize=size(population,1);
for i=1:popsize
    individual=population(i,:);
    for i=1:6
        flag=individual(i)-individual(
%-----

function y=targetfun(s); %目标函数
global data
global data_size

```

```

global flag
s=sort(s);
%解决相同元素问题
for i=2:6
    for j=2:i
        flag=s(i)-s(j-1);
        if flag==0
            s(i)=s(i)+(i-j+1);
        end
    end
end

s=sort(s);

Score = zeros( 1, data_size(1)/2 );
for i = 1 : data_size(1)/2
    Y = interp1( data( 2*i-1 , s(1, : ) ),data( 2*i,
s(1, : ) ),data( 2*i-1 ,2:51 ),'spline');
    dert_Y = abs( Y(:, :) - data( 2*i, 2:51 ) );

    for j = 1:( data_size(2) - 2 )
        if dert_Y( j ) <= 1
            Score( 1, i ) = Score( 1, i ) ;
        elseif dert_Y( j ) <= 2
            Score( 1, i ) = Score( 1, i ) + 1;
        elseif dert_Y( j ) <= 3
            Score( 1, i ) = Score( 1, i ) + 2;
        elseif dert_Y( j ) <= 5
            Score( 1, i ) = Score( 1, i ) + 4;
        else
            Score( 1, i ) = Score( 1, i ) + 10;
        end
    end
end
end
y=mean(Score(1,:));

```