

回归与内插



Notebook



MATLAB

<https://youtu.be/t5Hgmp4cFOA?si=d0zHmAm7MGFbSJml>

polynomial Curve Fitting:polyfit()

目的是根据给定区间或者区域上的有限个采样点的函数值，构造一个较简单的函数去逼近一个复杂或者未知的函数。



$p = \text{polyfit}(x,y,n)$ 返回次数为 n 的多项式 $p(x)$ 的系数，该阶数是 y 中数据的最佳拟合（基于最小二乘指标）。 p 中的系数按降幂排列， p 的长度为 $n+1$ 。

返回值 a 和 b ， a 为系数， b 为常数项， $y=ax+b$

- (1) $P=\text{polyfit}(X,Y,m)$
- (2) $[P,S]=\text{polyfit}(X,Y,m)$
- (3) $[P,S,\mu]=\text{polyfit}(X,Y,m)$

根据样本数据X和Y，产生一个m次多项式P及其在采样点误差数据S，mu是一个二元向量，mu(1)是平均数mean(X)，而mu(2)是std(X)。

Are x and y Linearly Correlated?

If not, the line may not well describe their relationship

Check the linearity by using:

scatter(): scatterplot

corrcoef(): correlation

coefficient, $-1 \leq r \leq 1$ ，越接近1，代表正相关相关性越强

Higher Order Polynomials

多项式项数越多，精度越高。

```
x = [-1.2 -0.5 0.3 0.9 1.8 2.6 3.0 3.5];
y = [-15.6 -8.5 2.2 4.5 6.6 8.2 8.9 10.0];
figure('Position', [50 50 1500 400]);
for i=4:6
    subplot(1,3,i-3);
    p = polyfit(x,y,i);
    xfit = x(1):0.1:x(end);
    yfit = polyval(p,xfit);
    plot(x,y,'ro',xfit,yfit);
    set(gca,'FontSize',14);
    ylim([-17, 11]);
    legend('Data points','Fitted curve');
end
```

多变量线性方程

```
regress()
```

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

$X = [\text{ones}(\text{length}(x_1), 1) \ x_1 \times 2]$

Interpolation 内插

Common Interpolation Approaches

- Piecewise linear interpolation
- Piecewise cubic polynomial interpolation
- Cubic spline interpolation

<code>interp1()</code>	1-D data interpolation (table lookup)
<code>pchip()</code>	Piecewise Cubic Hermite Interpolating Polynomial
<code>spline()</code>	Cubic spline data interpolation
<code>mkpp()</code>	Make piecewise polynomial

一次线性内插：interp1()



一维数值插值：被插值函数是一个单变量函数。

`Y1=interp1(X,Y,X1,'method')`

根据X,Y的值，计算函数在X1处的值。

method：插值方法：linear(默认),nearest,spline,cubic

其中线性插值方法的计算量与样本点n无关。n越大，误差越小。

linear：线性插值（默认）。将与插值点靠近的两个数据点用直线连接，然后在直线上选取对应插值点的数据。

nearest：最近点插值，选择最近样本点的值作为数据插值。如果最近的样本点在两个点中间，则取后一个样本点的值。

pchip：分段3次埃尔米特插值。采用分段三次多项式，除满足插值条件，还需满足在若干节点处相邻段插值函数的一阶导数相等，使得曲线光滑的同时，还具有保形性。（样本点的极值在插值以后还是极值）

spline：3次样条插值。每个分段内构造一个三次多项式，使其插值函数除满足插值条件外，还要求在各节点处具有连续的一阶和二阶导数。（比pchip进一步提高了曲线的光滑性）

比较平滑的内插：spline()

二维数值插值

```
Z1=interp2(X,Y,Z,X1,Y1,'method')
```

其中，X、Y是两个向量，表示两个参数的采样点，Z是采样点对应的函数值。X1、Y1是两个标量或向量，表示要插值的点。

method方法跟一维插值基本相同，但是

不支持pchip方法

三维数值插值：

```
V1=interp3(X,Y,Z,V,X1,Y1,Z1,'method')
```

X1的取值范围不能超出X的给定范围，否则，会给出“NaN”错误。

MATLAB中有一个专门的3次样条插值函数Y1=spline(X,Y,X1)，其功能及使用方法与函数Y1=interp1(X,Y,X1,'spline')完全相同。