第3章 曲线拟合的最小二乘法

纪庆革
软件工程与应用研究所，
数据科学与计算机学院，
中山大学，广州
Email: 1024180018@qq.com

在实际中，数据不可避免地会有误差，插值函数会将这些误差也包括在内。

因此，需要一种新的逼近原函数的手段：
①不要求过所有的点（可以消除误差影响）；
②尽可能表现数据的趋势，靠近这些点。

有时候， 问题本身不要求构造的函数过所有的点。如：5个风景点，要修一条公路S使得S为直线，且到所有风景点的距离和最小。

对如上2类问题，有一个共同的数学提法：找函数空间上的函数g，使得g到f的距离最小。

预备知识

定义1：**向量范数**

映射：$\|\cdot\|: R^n \to R^+ \cup \{0\}$ 满足：

①非负性 $\|X\| \geq 0$, 且$\|X\| = 0 \Leftrightarrow X = 0$

②齐次性 $\forall a \in R, \|aX\| = |a| \cdot \|X\|$

③三角不等式 $\|X+Y\| \leq \|X\| + \|Y\|$

称该映射为向量的一种**范数**

我们定义两点的**距离**为：$\|X-Y\|$

常见的范数有：$\|X\|_1 = \sum_{i=1}^{n} |x_i|, X = \{x_1, x_2, \Lambda, x_n\}$

$$\|X\|_2 = \sqrt{\sum_{i=1}^{n} (x_i)^2}, X = \{x_1, x_2, \Lambda, x_n\}$$

$$\|X\|_\infty = \max\{|x_i|\}, X = \{x_1, x_2, \Lambda, x_n\}$$

定义2：函数f，g的关于离散点列$\{x_i\}_{i=0}^{n}$ 的**离散内积**为：

$$(f, g)_D = \sum_{i=0}^{n} f(x_i) g(x_i)$$

定义3：函数f的**离散范数**为

$$\|f\|_D=\sqrt{\sum_{i=0}^{n} f(x_i)f(x_i)}$$

提示：该种内积，范数的定义与向量的2－范数一致

我们还可以定义函数的离散范数为：

$$\|f\|_D=\left\|\left(f(x_0),f(x_1),\cdots,f(x_n)\right)\right\|_\infty=\max\left|\left(f(x_0),f(x_1),\cdots,f(x_n)\right)\right|$$

$$\|f\|_D=\left\|\left(f(x_0),f(x_1),\cdots,f(x_n)\right)\right\|_1=\sum_{i=0}^{n}\left|f(x_i)\right|$$

曲线拟合的最小二乘问题

定义 $f(x)$为定义在区间**[a, b]**上的函数，$\{x_i\}_{i=0}^{n}$ 为区间上$n+1$个互不相同的点， Φ 为给定的某一函数类。求 Φ 上的函数 g(x)，满足 $f(x)$和 $g(x)$ 的距离最小

如果这种距离取为2－范数的话，称为最小二乘问题。

下面我们来看看最小二乘问题：

求 $g(x)\in\Phi$ 使得 $R_2=\sqrt{\sum_{i=0}^{n}\left(g(x_i)-f(x_i)\right)^2}$ 最小

设

$$\Phi=span\{\varphi_0,\varphi_1,\Lambda\ \varphi_n\}$$

$$g(x)=a_0\varphi_0(x)+\Lambda\ +a_n\varphi_n(x)$$

则 $$\|g(x)-f(x)\|_D=\min_{\varphi\in\Phi}\|\varphi(x)-f(x)\|_D$$

即 $$\|f(x)-\left(a_0\varphi_0(x)+\Lambda\ +a_n\varphi_n(x)\right)\|_D$$

关于系数 $\{a_0,a_1,\Lambda\ a_n\}$ 最小

$$\left\| f(x) - \left(a_0\varphi_0(x) + \cdots + a_n\varphi_n(x) \right) \right\|_D^2$$

$$= \| f \|_D^2 - 2(f, a_0\varphi_0(x) + \cdots + a_n\varphi_n(x))_D$$

$$+ \left\| a_0\varphi_0(x) + \cdots + a_n\varphi_n(x) \right\|_D^2$$

$$= \| f \|_D^2 - 2\sum_{k=0}^{n} a_k (f, \varphi_k) + \sum_{i,k=0}^{n} a_i a_k (\varphi_i, \varphi_k)$$

$$= Q(a_0, a_1, \cdots, a_n)$$

由于它关于系数 $\{a_0, a_1, \cdots a_n\}$ 最小，因此有：

$$\frac{\partial Q}{\partial a_i} = 0, i = 0, \cdots, n$$

即 $\displaystyle\sum_{k=0}^{n} a_k (\varphi_i, \varphi_k) = (f, \varphi_i), i = 0, \cdots, n$

写成矩阵形式有：

$$
\begin{pmatrix}
(\varphi_0,\varphi_0)_D & \Lambda & (\varphi_0,\varphi_n)_D \\
\mathrm{M} & \mathrm{O} & \mathrm{M} \\
(\varphi_n,\varphi_0)_D & \Lambda & (\varphi_n,\varphi_n)_D
\end{pmatrix}
\begin{pmatrix}
a_0 \\ \mathrm{M} \\ a_n
\end{pmatrix}
=
\begin{pmatrix}
(f,\varphi_0)_D \\ \mathrm{M} \\ (f,\varphi_n)_D
\end{pmatrix}
$$

由 $\{\varphi_0,\varphi_1,\Lambda\ \varphi_n\}$ 的线性无关性，知道该方程存在唯一解

法方程

例：① $y=a+bx$

第一步：函数空间的基 $\{1,x\}$ ，然后列出法方程

$$\begin{pmatrix} (1,1)_D & (1,x)_D \\ (x,1)_D & (x,x)_D \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}=\begin{pmatrix} (f,1)_D \\ (f,x)_D \end{pmatrix}$$

② $y=ax^2+b$

第一步：函数空间的基 $\{x^2,1\}$ ，然后列出法方程

$$\begin{pmatrix} (x^2,x^2)_D & (1,x^2)_D \\ (1,x^2)_D & (1,1)_D \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}=\begin{pmatrix} (f,x^2)_D \\ (f,1)_D \end{pmatrix}$$

x	-3	-2	-1	2	4
y	14.3	8.3	4.7	8.3	22.7

$$y = ax^2 + b$$

第一步：函数空间的基 $\{x^2, 1\}$ ，然后列出法方程

$$\begin{pmatrix} (x^2, x^2)_D & (1, x^2)_D \\ (1, x^2)_D & (1,1)_D \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} (f, x^2)_D \\ (f, 1)_D \end{pmatrix}$$

$$\begin{pmatrix} 370 & 34 \\ 34 & 5 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 563 \\ 58.3 \end{pmatrix} \quad \Rightarrow \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0.832716 \\ 7.49691 \end{pmatrix}$$

$$y=ae^{bx}$$

由 $\ln y=\ln a+bx$ ，可以先做 $y^*=a^*+bx$ $\quad y=e^{y^*}=e^{a^*}e^{bx}$

x	-3	-2	-1	2	4
y	14.3	8.3	4.7	8.3	22.7
$\ln y$	2.66026	2.11626	1.54756	2.11626	3.12236

$$\begin{pmatrix}(1,1)_D & (1,x)_D\\(1,x)_D & (x,x)_D\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}=\begin{pmatrix}(f,1)_D\\(f,x)_D\end{pmatrix}$$

$$\begin{pmatrix}5 & 0\\0 & 34\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}=\begin{pmatrix}11.5627\\2.9611\end{pmatrix}\Rightarrow\begin{pmatrix}a\\b\end{pmatrix}=\begin{pmatrix}2.31254\\0.0870912\end{pmatrix}$$

超定方程组的最小二乘解

给定线性代数方程组Ax=b, 其中

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, x=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

当m>n时， Ax=b为超定方程组。求x*使得 $\|Ax-b\|_2^2$ 取最小值。应用多元函数求极值方法，求出 $A^T Ax = A^T b$ 的解x*。x*被称为超定方程组Ax=b的最小二乘解。

谢谢！