计算方法及MATLAB实现

郑勋烨　编著

主审：　高世臣　褚宝增　王祖朝　王翠香

副审：　李少琪　李明霞　赵琳琳　赵俊芳

国防工业出版社

第 7 节　曲线拟合的最小二乘法

实际问题中经常需要通过观察一组或多组数据的变化规律，来预测这些数据所蕴涵的某种函数 $f(x)$（它对应着一条“待测”的曲线）的解析表达式．插值方法是可行的，但由于“基本插值条件”的限制，要求插值曲线精确地通过每个数据点，当有若干数据点事实上“离群索居” —- 偏离大多数数据所在的曲线时，我们使用插值方法获得的曲线可能会出现较大的波动，造成数值不稳定．

解决的方法是设法构造一条曲线，它能够反映大多数数据点变化的规律，而忽略个别“异端分子”造成的局部波动．这就是 曲线拟合 方法．显然，最简单的方法莫过于用一条直线 $S = a + bx$ 来近似“待测”的曲线 $f(x)$，即 直线拟合．此外，我们需要衡量直线与待测曲线的误差．倘若选取 $S(x)$ 与待测函数 $f(x)$ 在各数据点的 误差的平方和最小，则相应的方法称为 曲线的最小二乘拟合 (Least Square Method for Curve fitting)．二者结合起来就是 曲线的最小二乘直线拟合．

- 1. **曲线的最小二乘直线拟合** (Linear Fitting)

【**定义** 1. **曲线的最小二乘直线拟合** (Linear Fitting)】

设连续函数 $f(x) \in C[a,b]$ 在一组离散点集合

$x_1 < x_2 < \cdots\cdots < x_{m-1} < x_m$ 处给定

散点数据表$\{f(x_i) = y_i, i = 1, 2, \cdots, m\}$ 如下：

$$f(x_1) = y_1, f(x_2) = y_2, \cdots\cdots\cdots, f(x_m) = y_m$$

或者通常见到的表格形式：

x_i	x_1	x_2	x_3	$\cdots$	x_m
$f(x_i)$	$f(x_1)$	$f(x_2)$	$f(x_3)$	$\cdots$	$f(x_m)$

假设由两个幂函数 $1, x$ 构成的区间 $[a,b]$ 上的 线性无关连续函数系

$$\{\varphi_k(x) \in C[a,b] | k = 0,1\} = \{1, x\}$$

其 线性扩张函数空间 或 线性包 $S = span\{\varphi_k(x)\}$ 定义为由它们的线性组合形式作成的函数集合

$$S = span\{\varphi_k(x)\} = \{S(x) \in C[a,b] | S(x) = a_0 + a_1 x\}$$

就是区间 $[a,b]$ 上的 线性函数，几何上对应于直线.

现在要来构造一个函数 $S^*(x)$，满足条件:

(1)$S^*(x) \in span\{1, x\}$. 或存在系数 a_0^*, a_1^* 使得有线性表示

$$S^*(x) = a_0^* + a_1^* x$$

(2)$S^*(x)$ 与待测函数 $f(x)$ 在各数据点的
<u>绝对误差平方和最小</u>. 或

$$\sum_{i=1}^{m} |S^*(x_i) - f(x_i)|^2 = \sum_{i=1}^{m} |a_0^* + a_1^* x_i - f(x_i)|^2$$

$$= \min_{S(x)\in S} \sum_{i=1}^{m} |a_0 + a_1 x_i - f(x_i)|^2.$$

这就是 曲线的最小二乘直线拟合.

显然，适合用直线拟合的函数 $f(x)$，其数据分布应当近似呈现“线性关系”，几何上看就是大多数数据点 $(x_i, f(x_i))$(或写为 (x_i, y_i)) 分布在一条直线附近. 当然，在某些情况下，数据的非线性关系也能通过某种特定的变换转化为线性关系(参阅下文的例题).

如果标记绝对误差 $\delta_i = a_0 + a_1 x_i - f(x_i)$,绝对误差向量

$$\vec{\delta} = (\delta_1, \cdots, \delta_m)^T$$

则曲线的最小二乘直线拟合等价于寻找函数 $S^*(x)$ 使得绝对误差向量的欧氏范数最小:

$$\|\vec{\delta^*}\|_2^2 = \sum_{i=1}^{m} |a_0^* + a_1^* x_i - f(x_i)|^2$$

$$= \min_{S(x)\in S} \sum_{i=1}^{m} |a_0 + a_1 x_i - f(x_i)|^2 = \min_{S(x)\in S} \|\vec{\delta}\|_2^2.$$

这等价于求关于线性组合系数 a_0, a_1 的二元函数

$$F(a_0, a_1) = \sum_{i=1}^{m} (a_0 + a_1 x_i - f(x_i))^2$$

的极小值点 a_0^*, a_1^*.

或者标记 $f(x_i) = y_i$,$a_0 = a, a_1 = b$, 则曲线的最小二乘直线拟合等价于求二元函数

$$F(a, b) = \sum_{i=1}^{m} (a + b x_i - y_i)^2$$

的极小值点 a^*, b^*. 将 $F(a, b)$ 作变形得到两个容易求偏导数的形式

$$F(a,b) = \sum_{i=1}^{m}(a^2 + 2(bx_i - y_i)a + y_i^2)$$

$$= ma^2 + 2a\sum_{i=1}^{m}(bx_i - y_i) + \sum_{i=1}^{m} y_i^2.$$

以及

$$F(a,b) = \sum_{i=1}^{m}(x_i^2 b^2 + 2(a - y_i)x_i b + (a - y_i)^2)$$

$$= b^2\sum_{i=1}^{m} x_i^2 + 2b\sum_{i=1}^{m}(a - y_i)x_i + \sum_{i=1}^{m}(a - y_i)^2$$

由求多元函数极值的必要条件，我们有两个偏导数为零的条件 $\dfrac{\partial F}{\partial a} = 2ma + 2\sum_{i=1}^{m}(bx_i - y_i)$

$$= 2ma + 2b\sum_{i=1}^{m} x_i - 2\sum_{i=1}^{m} y_i = 0$$

以及

$$\frac{\partial F}{\partial b} = 2b\sum_{i=1}^{m} x_i^2 + 2\sum_{i=1}^{m}(a - y_i)x_i$$

$$= 2b\sum_{i=1}^{m} x_i^2 + 2a\sum_{i=1}^{m} x_i - 2\sum_{i=1}^{m} x_i y_i = 0$$

从而我们获得关于线性组合系数 a_0, a_1 的线性代数方程组

$$\left\{\begin{aligned} ma + b\sum_{i=1}^{m} x_i &= \sum_{i=1}^{m} y_i \\ a\sum_{i=1}^{m} x_i + b\sum_{i=1}^{m} x_i^2 &= \sum_{i=1}^{m} x_i y_i \end{aligned}\right. \qquad *$$

可称之为 最小二乘拟合直线正规方程组，简称 正规方程组或 法方程(Normal Equations).

由 Cramer 法则，其解为

$$\begin{cases} a = \dfrac{\displaystyle\sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} x_i y_i}{\displaystyle m\sum_{i=1}^{m} x_i^2 - (\sum_{i=1}^{m} x_i)^2} \\ b = \dfrac{\displaystyle m\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i}{\displaystyle m\sum_{i=1}^{m} x_i^2 - (\sum_{i=1}^{m} x_i)^2} \end{cases} \quad *$$

当然，这个解的表示公式不需要记忆．对于给定的数据 (x_i, y_i), 我们只要分别计算出 3 个系数

$$m, \quad \sum_{i=1}^{m} x_i, \quad \sum_{i=1}^{m} x_i^2$$

和 2 个非齐次项：

$$d_0 = \sum_{i=1}^{m} y_i, \qquad d_1 = \sum_{i=1}^{m} x_i y_i$$

则建立具体的 正规方程组 并直接求解就好了 (如果手边有计算器，可以使用 STAT 按钮统计功能)．显然，建立最小二乘拟合直线，主要的计算量在于“预处理”的过程：系数和非齐次项的计算上．为了明晰起见，可以列出一张数据表格来．

【注记 1. 内积表示与格拉姆矩阵】

倘若引入函数的某种“内积”标记

$$\langle 1,1\rangle := \sum_{i=1}^{m} 1 = m, \qquad \langle 1,x\rangle := \langle x,1\rangle = \sum_{i=1}^{m} x_i;$$

$$\langle x,x\rangle := \sum_{i=1}^{m} x_i^2$$

两个非齐次项亦可以定义为待测函数 $f(x)$ 与基函数 $x_i^k, k=0,1$ 的“内积”，形如：

$$d_k := \langle f, x_i^k \rangle = \sum_{i=1}^{m} f(x_i) x_i^k, \qquad k = 0, 1.$$

即

$$d_0 = \langle f, 1 \rangle = \sum_{i=1}^{m} f(x_i) = \sum_{i=1}^{m} y_i$$

$$d_1 = \langle f, x \rangle = \sum_{i=1}^{m} f(x_i) x_i = \sum_{i=1}^{m} x_i y_i$$

继而引入向量标记

$$\vec{a} = (a_0, a_1)^T = (a, b)^T, \qquad \vec{d} = (d_0, d_1)^T$$

则 正规方程组 可写为矩阵 - 向量形式的线性代数方程组

$$G\vec{a} = \vec{d}$$

即

$$\begin{pmatrix} m & \sum\limits_{i=1}^{m} x_i \\ \sum\limits_{i=1}^{m} x_i & \sum\limits_{i=1}^{m} x_i^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum\limits_{i=1}^{m} y_i \\ \sum\limits_{i=1}^{m} x_i y_i \end{pmatrix}$$

或即

$$\begin{pmatrix} \langle 1,1\rangle & \langle 1,x\rangle \\ \langle x,1\rangle & \langle x,x\rangle \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}=\begin{pmatrix} \langle f,1\rangle \\ \langle f,x\rangle \end{pmatrix}$$

其中 正规方程组 的 2 阶系数方阵称为 2 阶 格拉姆矩阵 (Gram Matrix)：

$$G = \begin{pmatrix} \langle 1,1\rangle & \langle 1,x\rangle \\ \langle x,1\rangle & \langle x,x\rangle \end{pmatrix} = \begin{pmatrix} m & \sum_{i=1}^{m} x_i \\ \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 \end{pmatrix}$$

显然，它是一个 对称矩阵 (Symmetric Matrix).

上述表示的更一般的形式可以自然推广到高维空间.

【注记 2. **最小二乘法** (Least Square Method) **的含义】**

关于误差大小的衡量，我们知道有多种方法，它们实际上对应于不同的 向量范数. 比如，标记各数据节点 (x_i, y_i) 的绝对误差为 $\delta_i = S_i - f(x_i)$, 绝对误差向量 为各节点的绝对误差 $\delta_i = S_i - f(x_i)$ 作为分量坐标构成的列向量

$$\vec{\delta} = (\delta_1, \cdots, \delta_m)^T$$

则曲线的最小二乘拟合等价于考虑绝对误差向量的
<u>欧氏范数的平方</u>

$$\|\vec{\delta}\|_2^2 = \sum_{i=1}^{m} |\delta_i|^2 = \sum_{i=1}^{m} |S_i - f(x_i)|^2$$

我们要寻找合适的拟合曲线 $S(x)$，让 <u>平方误差</u> 即绝对误差向量的 <u>欧氏范数的平方</u>$\|\vec{\delta}\|_2^2$ 达到最小值. 或者换言之，让 <u>均方误差</u> 即绝对误差向量的 <u>欧氏范数</u>$\|\vec{\delta}\|_2$ 达到最小值.

而其他的绝对误差向量的 范数 的选取方法比如有：

(1) 选取绝对误差向量的 最大模范数(即 $\infty-$ 范数):

$$\|\vec{\delta}\|_\infty := \max_{1\le i\le m} |\delta_i| = \max_{1\le i\le m} |S_i - f(x_i)|$$

(2) 选取绝对误差向量的 绝对和范数(即 $1-$ 范数):

$$\|\vec{\delta}\|_1 := \sum_{i=1}^{m} |\delta_i| = \sum_{i=1}^{m} |S_i - f(x_i)|$$

但上面这两种范数： $\|\vec{\delta}\|_\infty$ 和 $\|\vec{\delta}\|_1$ 均涉及绝对值的计算，处理起来不甚便利．而且不如 欧氏范数 那样脍炙人口：它直接对应于各待测曲线数据点和拟合曲线数据点的 欧氏距离 (平方和的平方根)．

因此我们最习惯采用欧氏范数来度量误差向量的大小．曲线的最小二乘拟合就是要构造一条曲线，使拟合曲线上面的点与待测曲线数据点的绝对误差向量的欧氏范数 $\|\vec{\delta}\|_2$ 最小．为了避免进行开方运算，又等价于让绝对误差向量的欧氏范数的平方 $\|\vec{\delta}\|_2^2$ 最小．这就是“最小二乘” (Least Squares) 的含义．

【释例】求 2 维列向量 $x=(\frac{1}{2},\frac{\sqrt{3}}{2})^T$ 的 3 种范数：

(1) 2- 范数 (欧氏范数)：

$$||x||_2=\sqrt{x_1^2+x_2^2}=\sqrt{\tfrac{1}{4}+\tfrac{3}{4}}=\sqrt{1}=1;$$

(2) ∞- 范数：

$$||x||_\infty=\max\{|x_1|,|x_2|\}=\max\{|\frac{1}{2}|,|\frac{\sqrt{3}}{2}|\}=\frac{\sqrt{3}}{2}\ \ ;$$

(3) 1- 范数 (绝对和范数)：

$$||x||_1 = \sum_{i=1}^{2} |x_i| = |\frac{1}{2}| + |\frac{\sqrt{3}}{2}| = \frac{1+\sqrt{3}}{2} \ .$$

我们看到，这里 ∞- 范数和 1- 范数之间有关系：

$$||x||_\infty = \frac{\sqrt{3}}{2} \leq ||x||_1 = \frac{1+\sqrt{3}}{2} \leq 2||x||_\infty = \sqrt{3}$$

【**例** 1. **曲线的最小二乘直线拟合** (Linear Fitting)】 炼钢是个氧化脱碳过程. 冶炼时间为函数 $f(x)$, 与含碳量 x 的多少有关. 设在一组离散点集合 $x_1 < x_2 < \cdots\cdots < x_5$ 处给定了如下散点数据表, 其中 x_i 为含碳量, $y_i = f(x_i)$ 为相应冶炼时间. 试求 $f(x)$ 的最小二乘拟合直线.

x_i	165	123	150	123	140	
y_i	187	126	172	125	148	

【 解 】

设最小二乘拟合直线方程为 $S(x) = a + bx$. 其 最小二乘拟合直线正规方程组 为

$$\begin{cases} ma + b\sum_{i=1}^{m} x_i = \sum_{i=1}^{m} y_i \\ a\sum_{i=1}^{m} x_i + b\sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} x_i y_i \end{cases} \quad *$$

简单计算各系数得 (如果手边有计算器，可以使用 STAT 按钮统计功能)

$$m = 5, \quad \sum_{i=1}^{m} x_i = 702,$$

$$\sum_{i=1}^{m} y_i = 758, \sum_{i=1}^{m} x_i^2 = 99864, \quad \sum_{i=1}^{m} x_i y_i = 108396$$

正规方程组即是

$$\begin{cases} 5a + 702b &= 758 \\ 702a + 99864b &= 108396 \end{cases} \qquad *$$

解得 $a \approx -60.9392, b \approx 1.5138$. 即最小二乘拟合直线方程为 $S(x) = -60.9392 + 1.5138x$.

【**例** 2. **曲线的最小二乘直线拟合** (Linear Fitting)】

观测质点的直线运动，获得位移 $s(t)$ 关于时间 t 的如下数据表．试求运动方程.

t_i	0	0.9	1.9	3	3.9	5
s_i	0	10	30	50	80	110

【 解 】

通过描点作图看到直线运动的位移 $s(t)$ 关于时间 t 的变化近似呈线性规律. 建立运动方程即构造 $s=s(t)$ 的最小二乘拟合直线.

设最小二乘拟合直线方程为 $S(t)=a+bt$. 其 最小二乘拟合直线正规方程组 为

$$
\left\{
\begin{aligned}
ma + b\sum_{i=1}^{m} t_i &= \sum_{i=1}^{m} s_i \\
a\sum_{i=1}^{m} t_i + b\sum_{i=1}^{m} t_i^2 &= \sum_{i=1}^{m} t_i s_i
\end{aligned}
\right. \qquad *
$$

即是

$$\begin{cases} 6a + 14.7b & = 280 \\ 14.7a + 53.63b & = 1078 \end{cases} \quad *$$

解得 $a \approx -7.855048, b \approx 22.253761$. 即最小二乘拟合直线运动方程为 $S(x) = -7.855048 + 22.253761t$. 且可得平方误差为绝对误差向量的 <u>欧氏范数的平方</u>

$$\|\vec{\delta}\|_2^2 = \sum_{i=1}^{6} |s_i - s(t_i)|^2 \approx 210$$

【**例** 3. **非线性模型转化为直线拟合问题** (NON-Linear Model)】

设在一组离散点集合 $x_1 < x_2 < \cdots\cdots < x_5$ 处给定了函数 $f(x)$ 的如下散点数据表. 试求形如 $y = ae^{bx}$ 的经验公式, 使得它与函数 $f(x)$ 拟合.

x_i	1.00	1.25	1.50	1.75	2.00
y_i	5.10	5.79	6.53	7.45	8.46

【解】

现在原始数据分布呈现的并非是线性关系，而是指数函数关系．我们需要首先将问题转化为最小二乘拟合直线的构造．指数函数经验公式方程为 $y = ae^{bx}$.取对数 得

$$S(x) = \ln y = \ln(ae^{bx}) = \ln a + \ln e^{bx} = \ln a + bx$$

即

$$S = A + bx, \qquad A = \ln a, a = e^{A}$$

线性关系散点数据表对应为

x_i	1.00	1.25	1.50	1.75	2.00
y_i	5.10	5.79	6.53	7.45	8.46
S_i	1.629	1.756	1.876	2.008	2.135

其 最小二乘拟合直线方程组(或称 法方程) 为

$$\begin{cases} mA + b\sum_{i=1}^{m} x_i = \sum_{i=1}^{m} S_i \\ A\sum_{i=1}^{m} x_i + b\sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} x_i S_i \end{cases} \quad *$$

即是

$$\begin{cases} 5A + 7.5b & = 9.404 \\ 7.5A + 11.875b & = 14.422 \end{cases} \quad *$$

解得 $A \approx 1.122, b \approx 0.5056. a = e^A = e^{1.122} = 3.071.$ 即最小二乘拟合直线方程为

$$S(x) = 1.122 + 0.505x$$

而指数函数经验公式为

$$y = ae^{bx} = 3.071e^{0.5056x}$$

【**例** 4. **非线性模型转化为直线拟合问题** (NON-Linear Model)】

设在某化学反应中获得分解物浓度 y 关于时间 t 的函数 $y = f(t)$ 在一组离散时刻点集合 $t_1 < t_2 < \cdots\cdots < t_{12}$ 处给定的的如下散点数据表. 描点发现函数近似呈现指数双曲关系. 试求形如 $y = ae^{-\frac{b}{t}}$, $a, b > 0$ 的经验公式, 使得它与浓度函数 $f(t)$ 拟合.

t_i	0	5	10	15	20	25
$y_i(\times 10^{-4})$	0	1.27	2.16	2.86	3.44	3.87

t_i	30	35	40	45	50	55
$y_i(\times 10^{-4})$	4.15	4.37	4.51	4.58	4.62	4.64

【 解 】

现在原始数据分布呈现的并非是线性关系，而是指数 + 双曲函数关系．我们需要首先将问题转化为最小二乘拟合直线的构造．指数函数经验公式方程为 $y = ae^{-\frac{b}{t}}$.取对数 得

$$S(x) = \ln y = \ln(ae^{-\frac{b}{t}}) = \ln a - \ln e^{\frac{b}{t}} = \ln a - \frac{b}{t}$$

即

$$S = A - bx, \qquad A = \ln a, x = \frac{1}{t}$$

抛弃初始时刻 $t=0$(此时反应没有开始，浓度自然为 0. 拟合条件自然满足. 从数学上解释，对于双曲函数 $\dfrac{1}{t}$ 此点无定义). 只留下 11 个数据点. 即对于 $m=11$, 其 最小二乘拟合直线方程组(或称 法方程) 为

$$\begin{cases} mA + b\sum_{i=1}^{m} x_i &= \sum_{i=1}^{m} S_i \\ A\sum_{i=1}^{m} x_i + b\sum_{i=1}^{m} x_i^2 &= \sum_{i=1}^{m} x_i S_i \end{cases} \qquad *$$

即是

$$\left\{\begin{aligned} 11A - 0.603975b &= -87.674095 \\ -0.603975A + 0.062321b &= 5.032489 \end{aligned}\right. \qquad *$$

解得 $A \approx -7.558781, b \approx 7.496163$. 于是 $a = e^A = e^{-7.558781} = 5.215103 \times 10^{-4}$. 即最小二乘拟合直线方程为

$$S(x) = -7.558781 \times 10^{-4} - 7.496163x$$

而指数函数经验公式为

$$y = ae^{-\frac{b}{t}} = 5.215103e^{-\frac{7.496163}{t}} \times 10^{-4}$$

2. 曲线拟合的一般问题

(General Curve fitting Problems)

将上述直线拟合的思路拓宽，我们选取的拟合直线能否改为拟合曲线？即用简单曲线来拟合一般的待测曲线？当然可以．最简单的函数依然是我们熟悉的多项式．方才我们选取的多项式空间基函数只有两个 $1, x$, 若再多取一个 x^2, 则对应的拟合曲线就是一个二次函数 (对应于抛物线)

$$S(x) = a + bx + cx^2$$

更一般地，基函数可取为 线性无关连续函数系

$\{\varphi_k(x) \in C[a,b] | k = 0, 1, 2, 3, \cdots n\} = \{\varphi_0(x), \varphi_1(x), \cdots, \varphi_n(x)\}$

然后寻求线性组合形式作成的函数

$$S(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \cdots + a_n\varphi_n(x)$$

来拟合一般的待测曲线. 这就是曲线的最小二乘拟合的一般方法.

【定义 1. 曲线的最小二乘拟合】 (Least Square Method for Curve fitting) 设连续函数 $f(x) \in C[a,b]$ 在一组离散点集合 $a = x_0 < x_1 < x_2 < \cdots\cdots < x_{m-1} < x_m = b$ 处给定
散点数据表$\{f(x_i) = y_i, i = 0, 1, 2, \cdots, m\}$ 如下：

$$f(x_0) = y_0, f(x_1) = y_1, f(x_2) = y_2, \cdots\cdots\cdots, f(x_m) = y_m$$

或者通常见到的表格形式： (注意：以下所有数据从 $i = 0$ 开始计数，总共有 $m + 1$ 个数据．下文相同.)

x_i	x_0	x_1	x_2	x_3	$\cdots$	x_m
$f(x_i)$	$f(x_0)$	$f(x_1)$	$f(x_2)$	$f(x_3)$	$\cdots$	$f(x_m)$

假设连续函数系

$\{\varphi_k(x) \in C[a,b] | k = 0,1,2,3,\cdots n\} = \{\varphi_0(x), \varphi_1(x), \cdots, \varphi_n(x)\}$

是区间 $[a,b]$ 上 线性无关函数系 ， 其 线性扩张函数空间 或 线性包 $S = span\{\varphi_k(x)\}$ 定义为由它们的线性组合形式作成的函数集合 $S = span\{\varphi_k(x)\}$

$= \{S(x) \in C[a,b] | S(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \cdots + a_n\varphi_n(x)\}$

【注记】 比如，若取 线性无关函数系 为

$$\{\varphi_k(x)\} = \{1, x, x^2, \cdots, x^n\}$$

则

$$S = span\{\varphi_k(x)\} = \{S(x)|S(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n\}$$

就是通常的 n 次代数多项式空间．其中的每个元素都是一个 n 次代数多项式 $S(x)$.

现在要来构造一个 线性包 $S = span\{\varphi_k(x)\}$ 中的函数 $S^*(x)$，满足条件:

(1)$S^*(x) \in span\{\varphi_k(x)\}$. 或存在系数 $a_0^*, a_1^*, \cdots, a_n^*$ 使得有线性表示

$$S^*(x) = a_0^*\varphi_0(x) + a_1^*\varphi_1(x) + \cdots + a_n^*\varphi_n(x)$$

(2)$S^*(x)$ 与待测函数 $f(x)$ 在各数据点的绝对误差平方和最小 (这等价于两个函数 $S^*(x)$ 与 $f(x)$ 的欧氏距离最小). 或

$$\sum_{i=0}^{m} |S^*(x_i) - f(x_i)|^2 =$$

$$\min_{S(x)\in S} \sum_{i=0}^{m} |S(x_i) - f(x_i)|^2$$

这就是 曲线的最小二乘拟合.

如果标记绝对误差 $\delta_i = S(x_i) - f(x_i)$,绝对误差向量

$$\vec{\delta} = (\delta_0, \delta_1, \cdots, \delta_m)^T$$

则曲线的最小二乘拟合等价于寻找函数 $S^*(x)$ 使得

绝对误差向量的欧氏范数最小: $\|\vec{\delta^*}\|_2^2 = \sum_{i=0}^{m} |S^*(x_i) - f(x_i)|^2$

$$= \min_{S(x)\in S} \sum_{i=0}^{m} |S(x_i) - f(x_i)|^2 = \min_{S(x)\in S} \|\vec{\delta}\|_2^2$$

由此，欲在所有 $S(x)\in span\{\varphi_k(x)\}$ 中求一函数 $S^*(x)$ ，即确定系数 $a_0^*, a_1^*, \cdots, a_n^*$ 使得

$$S^*(x)=\sum_{j=0}^{n} a_j^*\varphi_j(x)=a_0^*\varphi_0(x)+a_1^*\varphi_1(x)+\cdots+a_n^*\varphi_n(x)$$

满足

$$\|\vec{\delta^*}\|_2^2=\sum_{i=0}^{m}\omega(x_i)|S^*(x_i)-f(x_i)|^2=\min_{S(x)\in S}\|\vec{\delta}\|_2^2$$

由此我们获得 $n+1$ 阶非齐次线性代数方程组 $G\vec{a}=\vec{d}$ 如下：

$$\begin{pmatrix} (\varphi_0,\varphi_0) & (\varphi_0,\varphi_1) & \cdots & (\varphi_0,\varphi_n) \\ (\varphi_1,\varphi_0) & (\varphi_1,\varphi_1) & \cdots & (\varphi_1,\varphi_n) \\ \vdots & \vdots & \cdots & \vdots \\ (\varphi_n,\varphi_0) & (\varphi_n,\varphi_1) & \cdots & (\varphi_n,\varphi_n) \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \langle f,\varphi_0\rangle \\ \langle f,\varphi_1\rangle \\ \vdots \\ \langle f,\varphi_n\rangle \end{pmatrix}$$

这就是一般形式的 正规方程组 或 法方程(Normal Equations).

比如对于最小二乘二次拟合曲线，即曲线方程为 $S(x) = a + bx + cx^2$，相应 线性无关基函数系 由 $m+1=2+1=3$ 个幂函数构成：

$$\{\varphi_k(x)\} = \{1, x, x^2\}$$

引入函数的“内积”标记

$$\langle 1, 1\rangle := \sum_{i=0}^{m} 1 = m+1, \qquad \langle 1, x\rangle := \langle x, 1\rangle = \sum_{i=0}^{m} x_i;$$

$$\langle x, x\rangle := \sum_{i=0}^{m} x_i^2, \qquad \langle x, x^2\rangle := \sum_{i=0}^{m} x_i^3, \qquad \langle x^2, x^2\rangle := \sum_{i=0}^{m} x_i^4$$

两个非齐次项亦可以定义为待测函数 $f(x)$ 与基函数 $x_i^k, k=0,1,2$ 的“内积”，形如：

$$d_k := \langle f, x_i^k \rangle = \sum_{i=0}^{m} f(x_i) x_i^k, \qquad k = 0, 1.$$

即

$$d_0 = \langle f, 1 \rangle = \sum_{i=0}^{m} f(x_i) = \sum_{i=0}^{m} y_i$$

$$d_1 = \langle f, x \rangle = \sum_{i=0}^{m} f(x_i) x_i = \sum_{i=0}^{m} x_i y_i$$

$$d_2 = \langle f, x^2 \rangle = \sum_{i=0}^{m} f(x_i) x_i^2 = \sum_{i=0}^{m} x_i^2 y_i$$

于是 最小二乘二次拟合曲线方程组 为

$$\begin{pmatrix} m+1 & \sum_{i=0}^{m} x_i & \sum_{i=0}^{m} x_i^2 \\ \sum_{i=0}^{m} x_i & \sum_{i=0}^{m} x_i^2 & \sum_{i=0}^{m} x_i^3 \\ \sum_{i=0}^{m} x_i^2 & \sum_{i=0}^{m} x_i^3 & \sum_{i=0}^{m} x_i^4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{m} y_i \\ \sum_{i=0}^{m} x_i y_i \\ \sum_{i=0}^{m} x_i^2 y_i \end{pmatrix}$$

显然，由 Cramer 法则知道，非齐次线性代数方程组 $G\vec{a}=\vec{d}$ 存在唯一解 $\vec{a}^*$ 的充分必要条件是格拉姆系数矩阵非奇异.
但我们已知的

$$\{\varphi_k(x)\in C[a,b]|k=0,1,2,3,\cdots n\}=\{\varphi_0(x),\varphi_1(x),\cdots,\varphi_n(x)\}$$

在区间 $[a,b]$ 上 线性无关 ，这一条件并不能保证格拉姆矩阵 G 非奇异. 我们需要施加额外的限制，即所谓 哈尔条件 (Haar Condition).

【**定理** 1】 如果设连续函数系

$\{\varphi_k(x) \in C[a,b] | k = 0,1,2,3,\cdots n\} = \{\varphi_0(x), \varphi_1(x), \cdots, \varphi_n(x)\}$

在区间 $[a,b]$ 上的任意线性组合形式作成的函数

$$S(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \cdots + a_n\varphi_n(x)$$

在离散点集合

$a = x_0 < x_1 < x_2 < \cdots\cdots < x_{m-1} < x_m = b, \quad m \geq n$ 上 至多只有 n 个不同零点，即满足 哈尔条件 (Haar Condition).

则方程组 $G\vec{a}=\vec{d}$ 存在唯一解 $\vec{a}^*$, 并且以 $\vec{a}^*$ 的分量作为系数的线性组合形式作成的函数

$$S^*(x)=a_0^*\varphi_0(x)+a_1^*\varphi_1(x)+\cdots+a_n^*\varphi_n(x)$$

与待测函数 $f(x)$ 在各数据点的 绝对误差平方和最小：

$$\sum_{i=0}^{m}|S^*(x_i)-f(x_i)|^2=\min_{S(x)\in S}\sum_{i=0}^{m}|S(x_i)-f(x_i)|^2$$

即 $S^*(x)$ 就是 曲线的最小二乘拟合.

【例 1. 曲线的最小二乘二次拟合 (Quadric Fitting)】

设在一组离散点集合 $x_0 < x_1 < x_2 < \cdots\cdots < x_6$ 处给定了函数 $f(x) \in C[a,b]$ 的如下散点数据表. 试求 $f(x)$ 的最小二乘二次拟合曲线.

x_i	-3	-2	-1	0	1	2	3
y_i	1	0	0	0	0	1	2

【解】

设最小二乘二次拟合曲线方程为 $S(x)=a+bx+cx^2$. 其最小二乘二次拟合曲线方程组为

$$\begin{pmatrix} m+1 & \sum_{i=0}^{m} x_i & \sum_{i=0}^{m} x_i^2 \\ \sum_{i=0}^{m} x_i & \sum_{i=0}^{m} x_i^2 & \sum_{i=0}^{m} x_i^3 \\ \sum_{i=0}^{m} x_i^2 & \sum_{i=0}^{m} x_i^3 & \sum_{i=0}^{m} x_i^4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{m} y_i \\ \sum_{i=0}^{m} x_i y_i \\ \sum_{i=0}^{m} x_i^2 y_i \end{pmatrix}$$

本问题中 $m=6, m+1=7, n=2$. 简单计算得 (由于 Gram 矩阵的对称性，元素的实际计算量大为减少，几乎少了一半)

$$\sum_{i=0}^{6} x_i = -3-2-1+1+2+3=0$$

$$\sum_{i=0}^{6} x_i^2 = 9+4+1+1+4+9=28$$

$$\sum_{i=0}^{6} x_i^3 = -27 - 8 - 1 + 1 + 27 + 8 = 0$$

$$\sum_{i=0}^{6} x_i^4 = 81 + 16 + 1 + 1 + 16 + 81 = 196$$

$$\sum_{i=0}^{6} y_i = 1 + 1 + 2 = 4$$

$$\sum_{i=0}^{6} x_i y_i = -3 + 2 + 6 = 5$$

$$\sum_{i=0}^{6} x_i^2 y_i = 9 + 4 + 18 = 31$$

于是最小二乘二次拟合曲线方程组为

$$\begin{pmatrix} 7 & 0 & 28 \\ 0 & 28 & 0 \\ 28 & 0 & 196 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 31 \end{pmatrix}$$

解得 $a = -\frac{1}{7}, b = \frac{5}{28}, c = \frac{5}{28}$. 从而最小二乘二次拟合曲线为

$$y = -\frac{1}{7} + \frac{5}{28}x + \frac{5}{28}x^2$$