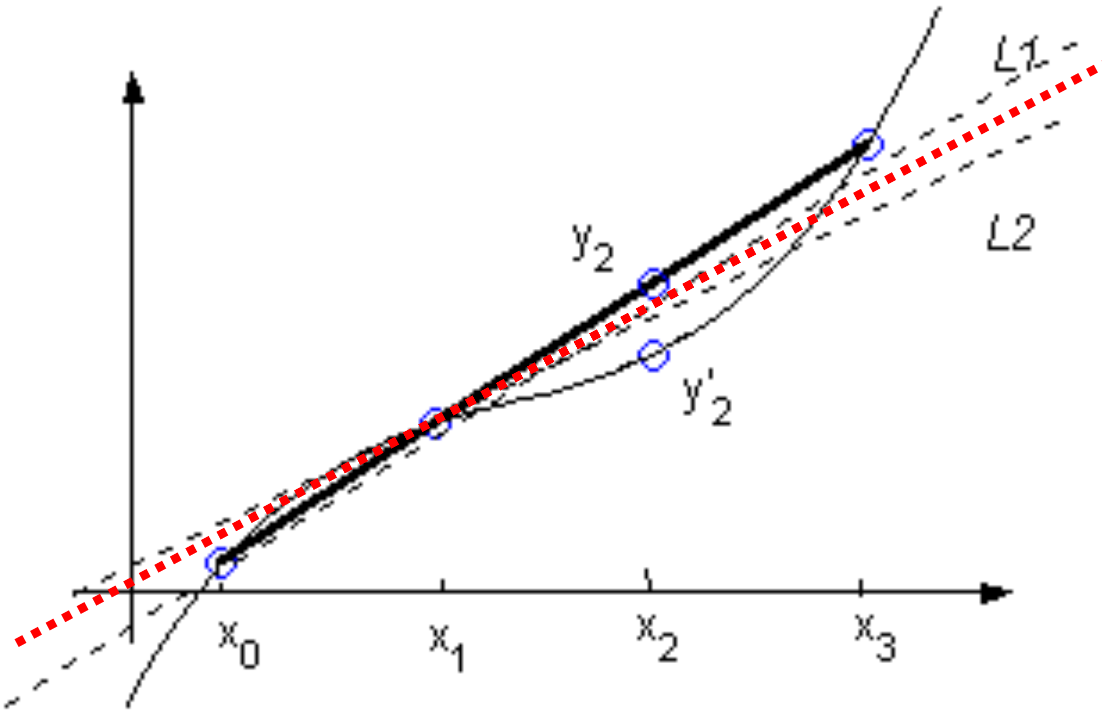
4.4 最小二乘拟合

数据建模

- 在实际的工程问题中，许多数据往往通过测试得到，这些数据不可避免的带有一定的误差。



No

- 增加节点约束来弱化由于“测不准”而造成的“差异”
- 解决方法：“满足节点约束”→“节点误差总体最小”拟合
- **插值：过点；（适合精确数据）**
- **拟合：不过点，整体近似；(有经验公式或有误差的数据)**

1. 最小二乘拟合定义

- **拟合问题**：设已知$x_1,\ldots,x_n$及 $y_i = f(x_i)$ $(i=1,2,\ldots,n)$, 要在一类曲线Φ中求一曲线$\varphi(x)$，使与$f(x)$在节点$x_1,\ldots,x_n$的误差$e_i = |y_i - \varphi(x_i)|$**总体**上最小。
- 上述e_i ($i=1,\ldots,n$)总体上最小，一般指误差向量$e=(e_1,\ldots,e_n)$的范数$||e||$最小。在拟合算法中一般取2-范数

$$\|e\|_2 = \sqrt{\sum_{i=1}^{n} e_i^2} = \sqrt{\sum_{i=1}^{n} [y_i - \varphi(x_i)]^2}$$

- 当误差向量取2范数时的拟合算法就是最小二乘拟合

例1　线性拟合

问题：对于给定的数据点 (x_i, y_i), $i = 1, 2, \cdots, n$，求拟合直线 $y = a + bx$，使总误差为最小，即在二元函数式中为最小。

$$\boldsymbol{Q}(a,b) = \sum_{i=1}^{N}[y_i - (a + bx_i)]^2$$

这里Q(a, b)是关于未知数a和b的二元函数，这一问题就是要确定a和b取何值时，二元函数

$$\boldsymbol{Q}(a,b) = \sum_{i=1}^{N}[y_i - (a + bx_i)]^2$$

的值最小?

线性拟合

由微积分的知识可知，这一问题的求解，可归结为求二元函数 $Q(a,b)$ 的极值问题，

即 a 和 b 应满足：

$$\frac{\partial Q}{\partial a}=0, \qquad \frac{\partial Q}{\partial b}=0,$$

$$\frac{\partial Q}{\partial a}=\sum_{i=1}^{N}2[y_i-(a+bx_i)]\cdot(-1)=0, \quad 即\ Na+b\sum_{i=1}^{N}x_i=\sum_{i=1}^{N}y_i,$$

$$\frac{\partial Q}{\partial b}=\sum_{i=1}^{N}2[y_i-(a+bx_i)]\cdot(-x_i)=0, 即\ a\sum_{i=1}^{N}x_i+b\sum_{i=1}^{N}x_i^2=\sum_{i=1}^{N}x_iy_i,$$

线性拟合

从而，未知数a，b 满足如下方程组(正则方程组)：

$$\begin{cases} \boldsymbol{N}a + b\sum_{i=1}^{\boldsymbol{N}} x_i = \sum_{i=1}^{\boldsymbol{N}} y_i\,, \\ a\sum_{i=1}^{\boldsymbol{N}} x_i + b\sum_{i=1}^{\boldsymbol{N}} x_i^2 = \sum_{i=1}^{\boldsymbol{N}} x_i\, y_i. \end{cases} \quad \textbf{(1)}$$

解出a，b，就可得最小二乘拟合直线 Q(a, b)=a+bx。

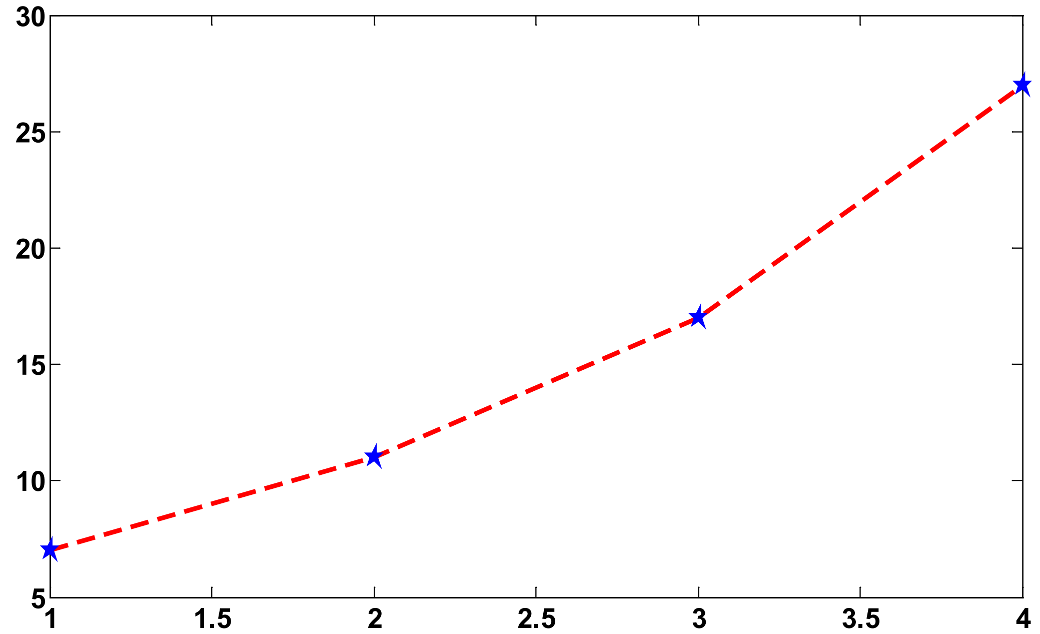
其它类型拟合问题

正如本节开头所指出的**最小二乘法并不只限于多项式，也可用于任何具体给出的函数形式**。特别重要的是有些非线性最小二乘拟合问题通过适当的变换可以转化为线性最小二乘问题求解。

线性化拟合

例2　已知数据表

x_i	1	2	3	4
y_i	7	11	17	27



求一形如 $y = Ae^{Bx}$ 的经验公式与已知数据拟合.

解：所求拟合函数是一个指数函数，对它两边取自然对数，得

$$\ln y = \ln A + Bx$$

线性化拟合

于是对应于上述数据表得到另一个数据表：

x_i	1	2	3	4
$\ln y_i$	1.95	2.40	2.83	3.30

x_i	1	2	3	4
z_i	1.95	2.40	2.83	3.30

若记 $z = \ln y, a_0 = \ln A, a_1 = B$ 则

$$z = a_0 + a_1 x$$

从而将原问题转化为由新数据表所给出的线性拟合问题。易知其正则方程组为：

线性化拟合

$$\begin{cases} \boldsymbol{N}a_0 + a_1\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} \boldsymbol{z}_i, \\ a_0\sum_{i=1}^{N} x_i + a_1\sum_{i=1}^{N} x_i^2 = \sum_{i=1}^{N} x_i\,\boldsymbol{z}_i. \end{cases} \quad \textbf{(1)}$$

由公式（1）得， $\begin{cases} 4a_0 + 10a_1 = 10.48, \\ 10a_0 + 3a_1 = 28.44, \end{cases}$

解得 $a_0 = 1.50, a_1 = 0.448$

于是 $\ln y = 1.50 + 0.448x$

故所求经验公式为

$$y = e^{1.50+0.448x} = 4.48e^{0.448x}$$

超定方程组（矛盾方程组）

试求下列超定方程组的解：

$$\begin{cases} x-15.5=0, \\ y-6.1=0, \\ x+y=20.9. \end{cases}$$

很显然，直接求解是不行的，因为满足方程组的精确解是不存在的！只能求出尽量满足方程组的近似解。

超定方程组

运用最小二乘法，要求满足方程组的解，即求使各方程误差平方和 u 最小的解 x, y, 就是方程组的近似解：

$$u = (x-15.5)^2 + (y-6.1)^2 + (x+y-20.9)^2$$

$$\begin{cases} \dfrac{\partial u}{\partial x} = 2(x-15.5) + 2(x+y-20.9) = 0, \\ \dfrac{\partial u}{\partial y} = 2(y-6.1) + 2(x+y-20.9) = 0. \end{cases}$$

得近似解：

$$\begin{cases} x = 15.26667, \\ y = 5.86667. \end{cases}$$

2. 最小二乘拟合法的一般理论：

内积和范数

- 函数$f(x)$和$g(x)$ 在节点上x_1、x_2、… x_n上取值向量的内积:

$$(f,g)=\sum_{i=1}^{n} f(x_i)g(x_i)=[\boldsymbol{f}(x_1),\boldsymbol{f}(x_2),\ldots\boldsymbol{f}(x_n)]\begin{bmatrix}\boldsymbol{g}(\boldsymbol{x}_1)\\ \boldsymbol{g}(\boldsymbol{x}_2)\\ \vdots\\ \boldsymbol{g}(\boldsymbol{x}_n)\end{bmatrix}$$

- 函数$f(x)$ 在节点上x_1、x_2、… x_n上取值向量的2-范数:

$$\|f\|=\sqrt{(f,f)}=\sqrt{[\boldsymbol{f}(x_1),\boldsymbol{f}(x_2),\ldots\boldsymbol{f}(x_n)]\begin{bmatrix}\boldsymbol{f}(\boldsymbol{x}_1)\\ \boldsymbol{f}(\boldsymbol{x}_2)\\ \vdots\\ \boldsymbol{f}(\boldsymbol{x}_n)\end{bmatrix}}$$

- 对任意函数f, g, h和数λ，有

  (i) $(f, g+h) = (f, g)+(f, h)$；

  (ii) $(\lambda f, g) = \lambda(f, g)$;

  (iii) $(f, g) = (g, f)$;

  (iv) $\|f\| \geq 0$且等号成立$\Leftrightarrow f=0$. （定义2.1，正定性）

- **注意:内积和范数依赖于节点选取.**

基函数

称函数$\varphi_0(x), \varphi_1(x), \ldots, \varphi_m(x)$关于节点$x_1,\ldots,x_n$ 线性无关，如果

它们的取值向量$\begin{pmatrix}\varphi_0(x_1)\\ \varphi_0(x_2)\\ \vdots\\ \varphi_0(x_n)\end{pmatrix}, \begin{pmatrix}\varphi_1(x_1)\\ \varphi_1(x_2)\\ \vdots\\ \varphi_1(x_n)\end{pmatrix}, \ldots, \begin{pmatrix}\varphi_m(x_1)\\ \varphi_m(x_2)\\ \vdots\\ \varphi_m(x_n)\end{pmatrix}$线性无关,

即只有当k_0，k_1，$\ldots,k_m$全为零时$k_0\varphi_0(x_i)+k_1\varphi_1(x_i)+\ldots+k_m\varphi_m(x_i)=0$, $i=1,2,\ldots,n$才全部成立

函数空间

• 线性无关函数φ_0，φ_1，…，φ_m的线性组合全体Φ称为由φ_0，φ_1，…，φ_m张成的函数空间(functional space)，记

$$\Phi = \mathrm{span}\{\varphi_0, \varphi_1, \ldots, \varphi_m\}$$

$$=\{\varphi(x)=a_0\varphi_0(x)+a_1\varphi_1(x)+\ldots+a_m\varphi_m(x) \mid a_0, a_1, \ldots, a_m \in R\},$$

则$\varphi_0, \varphi_1, \ldots, \varphi_m$称为$\Phi$的基函数。

最小二乘拟合

- 最小二乘拟合用数学语言表达：已知数据x_i, $y_i=f(x_i)$ (i=1,…,n) 和函数空间Φ=span{φ_0, φ_1, … φ_m }, 求函数$\varphi^*\in\Phi$，使

$$\|f-\varphi^*\| = \min \|f-\varphi\| .$$

- 问题等价于: 求a^*_0, a^*_1, …,$a^*_m\in R$，使

$$S(a^*_0, a^*_1, \ldots, a^*_m) = \min S(a_0, a_1, \ldots, a_m).$$

其中 $S(a_0, a_1, \ldots, a_m) = \|f-\varphi\|^2 = \sum_{i=1}^{n}[y_i - \sum_{j=0}^{m} a_j\,\varphi_j(x_i)]^2$

- $\varphi(x)=a_0\varphi_0(x)+a_1\varphi_1(x)+\ldots+a_m\varphi_m(x)=\sum_{j=0}^{m} a_j\,\varphi_j(x_i)$

(二次函数最小化)

函数极值的必要条件

- 对S求关于$a_0, a_1, \ldots, a_m$ 的偏导，

$$\frac{\partial S}{\partial a_k} = 0, k = 0,1,\cdots,m$$

$$-2\sum_{i=1}^{n}[y_i - \sum_{j=0}^{m} a_j \varphi_j(x_i)]\,\varphi_k(x_i) = 0$$

$$\boxed{\sum_{j=0}^{m}\sum_{i=1}^{n} a_j \varphi_j(x_i)\,\varphi_k(x_i) = \sum_{i=1}^{n} y_i \varphi_k(x_i) = \sum_{i=1}^{n} f(x_i)\,\varphi_k(x_i)}$$

表示： $\sum_{j=0}^{m}(\varphi_j, \varphi_k)a_j = \sum_{j=0}^{m} a_j\varphi_j(x_1)\,\varphi_k(x_1) + \cdots + a_j\varphi_j(x_n)\varphi_k(x_n)$

$= (f, \varphi_k), k = 0, \ldots, m$ （线性方程组）

$$\begin{cases} (\boldsymbol{\varphi}_0,\boldsymbol{\varphi}_0)\boldsymbol{a_0}+(\boldsymbol{\varphi_1},\boldsymbol{\varphi}_0)\boldsymbol{a_1}+\cdots+(\boldsymbol{\varphi_m},\boldsymbol{\varphi}_0)\boldsymbol{a_m}=(\boldsymbol{f},\boldsymbol{\varphi}_0) \\ (\boldsymbol{\varphi}_0,\boldsymbol{\varphi_1})\boldsymbol{a_0}+(\boldsymbol{\varphi_1},\boldsymbol{\varphi_1})\boldsymbol{a_1}+\cdots+(\boldsymbol{\varphi_m},\boldsymbol{\varphi_1})\boldsymbol{a_m}=(\boldsymbol{f},\boldsymbol{\varphi_1}) \\ \vdots \\ (\boldsymbol{\varphi}_0,\boldsymbol{\varphi_m})\boldsymbol{a_0}+(\boldsymbol{\varphi_1},\boldsymbol{\varphi_m})\boldsymbol{a_1}+\cdots+(\boldsymbol{\varphi_m},\boldsymbol{\varphi_m})\boldsymbol{a_m}=(\boldsymbol{f},\boldsymbol{\varphi_m}) \end{cases}$$

- 其矩阵形式为： <u>法方程组(正规方程组)</u>

$$\begin{pmatrix} (\varphi_0,\varphi_0) & (\varphi_0,\varphi_1) & \cdots & (\varphi_0,\varphi_m) \\ (\varphi_1,\varphi_0) & (\varphi_1,\varphi_1) & \cdots & (\varphi_1,\varphi_m) \\ \cdots & \cdots & \cdots & \cdots \\ (\varphi_m,\varphi_0) & (\varphi_m,\varphi_1) & \cdots & (\varphi_m,\varphi_m) \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} (f,\varphi_0) \\ (f,\varphi_1) \\ \vdots \\ (f,\varphi_m) \end{pmatrix} \quad (*)$$

法方程组

- **定理4.3** 如果函数$\varphi_0(x), \varphi_1(x), \ldots, \varphi_m(x)$关于节点$x_1, \ldots, x_n$线性无关，则

  1）法方程组(*)的解存在唯一；

  2）法方程组(*)的解是最小二乘拟合的唯一最优解。

证明略，流程图如4-4

例题 4.10（P91）

- 已知$\sin 0 = 0, \sin\frac{\pi}{6} = \frac{1}{2}, \sin\frac{\pi}{3} = \frac{\sqrt{3}}{2}, \sin\frac{\pi}{2} = 1$
- 由最小二乘法求$\sin x$的拟合曲线$\varphi(x) = ax + bx^3$
- 解题思路：
  - 确定$\varphi_0(x) = x, \varphi_1(x) = x^3, f(x) = \sin(x)$
  - 计算出$(\varphi_j, \varphi_k), (f, \varphi_k), k = 0,1$得到法方程组

    解得系数 a，b
  - 构建最小二乘拟合曲线$\varphi(x) = ax + bx^3$

3. 正交最小二乘拟合

多项式拟合

- 基函数一般取幂函数$\varphi_0(x)=1$, $\varphi_1(x)=x$, …, $\varphi_m(x)=x^m$ ，法方程组为:

$$\begin{pmatrix} n & \sum_{i=1}^{n} x_i & \cdots & \sum_{i=1}^{n} x_i^m \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \cdots & \sum_{i=1}^{n} x_i^{m+1} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{i=1}^{n} x_i^m & \sum_{i=1}^{n} x_i^{m+1} & \cdots & \sum_{i=1}^{n} x_i^{2m} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \\ \vdots \\ \sum_{i=1}^{n} x_i^m y_i \end{pmatrix}$$

高阶多项式拟合的病态

令$A=\begin{pmatrix}1 & x_1 & x_1^2 & \cdots & x_1^m\\ 1 & x_2 & x_2^2 & \cdots & x_2^m\\ \vdots & \vdots & \vdots & & \vdots\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_n & x_n^2 & \cdots & x_n^m\end{pmatrix}$,则上式化为 $A^TA\begin{pmatrix}a_0\\ a_1\\ \vdots\\ a_m\end{pmatrix}=A^T\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}$.

- ***m***很大时, 法方程组病态
- 系数矩阵**A^TA**的条件数太大!

- 避免病态: 选取函数类Φ的正交基函数$\psi_0, \psi_1, \ldots, \psi_m$,法方程组就成为简单的对角方程组, 其解

$$a_k=\frac{(f,\psi_k)}{(\psi_k,\psi_k)}, k=0,1,\cdots,m$$

$$\psi_m=\varphi_m-\frac{(\varphi_m,\psi_0)}{(\psi_0,\psi_0)}\psi_0(x)-\frac{(\varphi_m,\psi_1)}{(\psi_1,\psi_1)}\psi_1(x)-\cdots-\frac{(\varphi_m,\psi_{m-1})}{(\psi_{m-1},\psi_{m-1})}\psi_{m-1}(x)$$

正交最小二乘拟合

定义：给定节点$x_1, \ldots, x_n$和函数f与g,如果（f，g）$= 0$,称f与g关于节点$x_1, \ldots, x_n$**正交**。如果函数类Φ的基函数ψ_0、$\psi_1 \ldots \varphi_m$两两正交，则称为**一组正交基**。

- **Gram-Schmit正交化方法**

$$\psi_0(x) = \varphi_0(x), \qquad \psi_1(x) = \varphi_1(x) - \frac{(\varphi_1, \psi_0)}{(\psi_0, \psi_0)}\psi_0(x),$$

$$\psi_m(x) = \varphi_m(x) - \frac{(\varphi_m, \psi_0)}{(\psi_0, \psi_0)}\psi_0(x) - \frac{(\varphi_m, \psi_1)}{(\psi_1, \psi_1)}\psi_1(x), \ldots$$
$$- \frac{(\varphi_m, \psi_{m-1})}{(\psi_{m-1}, \psi_{m-1})}\psi_{m-1}(x)$$

例4.11

解法1：一般的多项式拟合

解法2：先进行施密特正交化，再解对角方程组

解法3：为了减轻病态问题的影响，采用递推正交化公式再求解

$\psi_0(x) = 1,$

$\psi_1(x) = (x - \alpha_0)\psi_0(x), \alpha_0 = \dfrac{(x\psi_0, \psi_0)}{(\psi_0, \psi_0)}$

$\psi_{k+1}(x) = (x - \alpha_k)\psi_k(x) - \beta_k \psi_{k-1}(x),$

$\alpha_k = \dfrac{(x\psi_k, \psi_k)}{(\psi_k, \psi_k)}, \beta_k = \dfrac{(\psi_k, \psi_k)}{(\psi_{k-1}, \psi_{k-1})}, k = 1, .. m - 1$

16. 最小二乘法解超定线性方程组

$$\begin{cases} 4x_1 + 2x_2 = 2 \\ 3x_1 - x_2 = 10 \\ 11x_1 + 3x_2 = 8 \end{cases}$$

解：$g(x_1, x_2) = (4x_1 + 2x_2 - 2)^2 + (3x_1 - x_2 - 10)^2 + (11x_1 + 3x_2 - 8)^2$

$= 16x_1^2 + 4x_2^2 + 4 + 16x_1x_2 - 16x_1 - 8x_2 + 9x_1^2 + x_2^2 + 100 - 6x_1x_2 - 60x_1 + 20x_2 + 121x_1^2 + 9x_2^2 + 64 + 66x_1x_2 - 176x_1 - 48x_2$

$= 146x_1^2 + 14x_2^2 + 73x_1x_2 + 168 - 252x_1 - 36x_2$

$$\frac{\partial g(x_1, x_2)}{\partial x_1} = 292x_1 + 73x_2 - 252 = 0$$

$$\frac{\partial g(x_1, x_2)}{\partial x_2} = 28x_2 + 73x_1 - 36 = 0$$

17. 求 $f(x) = x^4 + 3x^2 - 1$ 关于节点 $-1, 0, 1, 2, 3$ 的最小二乘三次逼近多项式：

(1) 基 $1, x, x^2, x^3$

(2) 基 $1, x, x^2, x^3$ 正交化

解：$n = 5$ $f(-1) = 3$ $f(0) = -1$ $f(1) = 3$ $f(2) = 27$ $f(3) = 81 + 27 - 1 = 107$

$\sum_{i=1}^{5} x_i = 5$

$\sum_{i=1}^{5} x_i^2 = 1 + 1 + 4 + 9 = 15$

$\sum_{i=1}^{5} x_i^3 = 8 + 27 = 35$

$\sum_{i=1}^{5} x_i^4 = 1 + 1 + 16 + 81 = 99$

$\sum_{i=1}^{5} x_i^5 = 32 + 243 = 275$

$\sum_{i=1}^{5} x_i^6 = 64 + 729 = 793$

$$\psi_m(x) = \varphi_m(x) - \frac{(\varphi_m, \psi_0)}{(\psi_0, \psi_0)}\psi_0(x) - \frac{(\varphi_m, \psi_1)}{(\varphi_m, \psi_1)}\psi_1(x) - \cdots - \frac{(\varphi_m, \psi_{m-1})}{(\psi_{m-1}, \psi_{m-1})}\psi_{m-1}(x)$$

$$a_k = \frac{(f, \psi_k)}{(\psi_k, \psi_k)}$$