



实验物理中的统计方法

第八章：矩方法

杨振伟

回顾

1. 最小二乘法与最大似然法的联系

对于高斯变量 y_i , 二者相同

2. 线性的最小二乘法估计

通过求矩阵的逆完成估计, 估计量是测量量 y_i 的线性函数

3. 非线性的最小二乘法估计

通过迭代完成估计, 方差可采用线性情况 $\chi^2 = \chi_{min}^2 + 1$ 来估计

4. 约束条件下的最小二乘法拟合

在约束条件下引入拉格朗日乘子改进实验观测量的精度

5. 用最小二乘法检验拟合优度

用 χ_{min}^2 作拟合优度统计, 满足 $N - m$ 自由度下的卡方分布

6. 用最小二乘法处理分区数据

把 y_i 当作泊松变量, 不确定用 λ_i 估计, 或用 y_i 估计 (推广最小二乘法)

7. 不等精度关联实验结果的并合问题

对存在相关性的数据的处理, 不确定度的修正

本章要点

- 矩的定义
- 矩方法估计量及其应用
- 矩方法与最大似然法和最小二乘法的比较

简单矩、原点矩和中心矩

假设随机变量 $x \sim f(x)$ 。

定义围绕某固定值 x_0 的 k 阶矩或简单矩 μ_k 为

$$m_k \equiv \int (x - x_0)^k f(x) dx$$

$x_0 = 0 \rightarrow k$ 阶原点矩 μ'_k 或 e_k :

$$e_k = \mu'_k \equiv \int x^k f(x) dx$$

一阶原点矩就是随机变量 x 的均值或数学期待值:

$$E[x] = \int x f(x) dx = \mu = \mu'$$

$x_0 = E[x] = \mu \rightarrow k$ 阶中心矩 μ_k :

$$\mu_k = \int (x - \mu)^k f(x) dx$$

二阶中心矩就是随机变量 x 的方差:

$$V[x] = \int (x - \mu)^2 f(x) dx = \sigma^2 = \mu_2$$

代数矩与中心矩的关系

原点矩

$$\begin{aligned}\mu'_0 &= 1 \\ \mu'_1 &= \mu \\ \mu'_2 &= \sigma^2 + \mu^2\end{aligned}$$

中心矩

$$\begin{aligned}\mu_0 &= 1 \\ \mu_1 &= 0 \\ \mu_2 &= \sigma^2\end{aligned}$$

低阶矩之间的关系

$$\mu_2 = \mu'_2 - \mu^2$$

一般情况下，它们的关系可以表示如下

$$\begin{aligned}\mu_k &= \sum_{l=0}^k \binom{k}{l} \mu'_{k-l} (-\mu'_1)^l \\ \mu'_k &= \sum_{l=0}^k \binom{k}{l} \mu_{k-l} (\mu'_1)^l\end{aligned}$$

□ 高阶矩对研究概率密度在大 $|x - \mu|$ 区间的行为很有帮助。

□ 对称分布的所有奇数中心矩为零。

矩的一般表达式

假设随机变量 $x \sim f(x; \vec{\theta})$, $\vec{\theta} = (\theta_1, \dots, \theta_m)$ 为未知参数。

目标：利用 n 个测量值 x_1, \dots, x_n 估计 $\vec{\theta}$ 。

构造 m 个线性独立函数 $a_i(x)$, $i = 1, \dots, m$, 其均值可写为

$$E[a_i(x)] = \int a_i(x) f(x; \vec{\theta}) dx \equiv e_i(\vec{\theta}) \quad \text{矩的一般表达式}$$

$e_i(\vec{\theta})$ 可以用无偏的样本均值来估计：

$$\hat{e}_i = \bar{a}_i = \frac{1}{n} \sum_{j=1}^n a_i(x_j)$$

令均值 $e_i(\vec{\theta})$ 等于样本均值 \bar{a}_i , 即 $e_i(\vec{\theta}) = \bar{a}_i = \hat{e}_i$, 那么

参数 $\vec{\theta}$ 的估计值 $\hat{\vec{\theta}}$ 可以通过求解 m 个 $e_i(\vec{\theta})$ 方程来确定。

矩方法估计量

求解 m 个 $e_i(\vec{\theta})$ 构成的方程组确定参数 $\vec{\theta}$ 的估计量 $\hat{\vec{\theta}}$:

$$\begin{aligned} e_1(\vec{\theta}) &= \frac{1}{n} \sum_{j=1}^n a_1(x_j) = \hat{e}_1 \\ &\vdots \\ e_m(\vec{\theta}) &= \frac{1}{n} \sum_{j=1}^n a_m(x_j) = \hat{e}_m \end{aligned}$$



$$\hat{\theta}_1 = \hat{\theta}_1(\hat{e}_1, \dots, \hat{e}_m)$$

$$\vdots$$

$$\hat{\theta}_m = \hat{\theta}_m(\hat{e}_1, \dots, \hat{e}_m)$$

$$\text{矢量形式: } \hat{\vec{\theta}} = \hat{\vec{\theta}}(\hat{\vec{e}})$$

$a_i(x)$ 原则上可以任意选择, 只要其期待值可计算并构成 m 个线性独立的函数。

常见的一种选择是: $a_i = x^i$

$$E[a_i(x)] = E[x^i]: x \text{ 的 } i \text{ 阶原点矩}$$

矩方法名称的来源

$\hat{\vec{\theta}}$: 矩方法估计量

矩方法估计量的协方差

我们还希望了解矩方法估计量 $\hat{\theta}$ 的协方差。

$$\begin{aligned}\hat{\theta}_1 &= \hat{\theta}_1(\hat{e}_1, \dots, \hat{e}_m) \\ &\vdots \\ \hat{\theta}_m &= \hat{\theta}_m(\hat{e}_1, \dots, \hat{e}_m)\end{aligned}$$



$$\text{cov}[\hat{\theta}_i, \hat{\theta}_j] = \sum_{k,l=1}^m \frac{\partial \hat{\theta}_i}{\partial \hat{e}_k} \frac{\partial \hat{\theta}_j}{\partial \hat{e}_l} \text{cov}[\hat{e}_k, \hat{e}_l]$$

$$\begin{aligned}\text{cov}[\hat{e}_k, \hat{e}_l] &= \text{cov}[\bar{a}_k, \bar{a}_l] = \text{cov}\left[\frac{1}{n} \sum_{i=1}^n a_k(x_i), \frac{1}{n} \sum_{j=1}^n a_l(x_j)\right] \\ &= \frac{1}{n^2} \sum_{i,j=1}^n \text{cov}[a_k(x_i), a_l(x_j)] \\ &= \frac{1}{n^2} \left(\sum_{i=j=1}^n \text{cov}[a_k(x_i), a_l(x_j)] + \sum_{i \neq j=1}^n \text{cov}[a_k(x_i), a_l(x_j)] \right) \\ &= \frac{1}{n} \text{cov}[a_k, a_l] \quad x_i \text{ 和 } x_j \text{ 相互独立} \rightarrow a_k(x_i) \text{ 和 } a_l(x_j) \text{ 相互独立}\end{aligned}$$

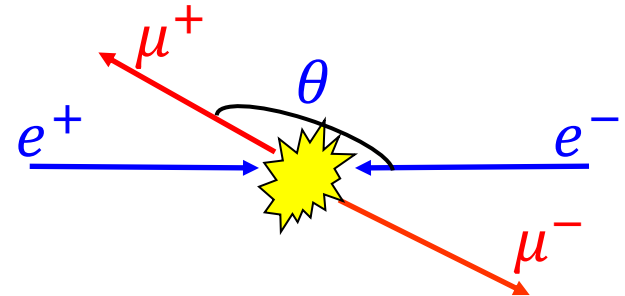
$\text{cov}[a_k, a_l]$ 可以用样本协方差估计

$$\widehat{\text{cov}}[a_k, a_l] = \frac{1}{n-1} \sum_{i=1}^n (a_k(x_i) - \bar{a}_k)(a_l(x_i) - \bar{a}_l)$$

例：角分布理论的简单验证

在 $e^+e^- \rightarrow \mu^+\mu^-$ 实验中，理论预言角分布为

$$\frac{dn}{d\cos\theta} \propto n(1 + \cos^2\theta) \quad n = \text{事例数}$$



$$E[\cos^2\theta] = \frac{\int_{-1}^{+1} \cos^2\theta \cdot n(1 + \cos^2\theta) d\cos\theta}{\int_{-1}^{+1} n(1 + \cos^2\theta) d\cos\theta} = 0.4 \quad \text{二阶原点矩}$$

为了验证理论，计算样本的二阶原点矩：

$$\overline{\cos^2\theta} = \frac{1}{n} \sum_{i=1}^n \cos^2\theta_i$$

假设检验可通过比较二阶原点矩与二阶样本原点矩来完成。

角分布理论的简单验证：不确定度估计

不含参数，已知 $\cos^2 \theta$ 的真值：

$$E[\cos^2 \theta] = 0.4$$

$\cos^2 \theta$ 的样本方差：

$$S^2 = \frac{1}{n} \sum_{i=1}^n (\cos^2 \theta_i - E[\cos^2 \theta])^2$$

样本的二阶矩 $\overline{\cos^2 \theta}$ 的方差：

$$V[\overline{\cos^2 \theta}] = \frac{S^2}{n}$$

假设实验观测1000次， $\overline{\cos^2 \theta} = 0.39$ ，并计算得到
 $S^2 = 0.15$ 。则实验结果报告为： 0.39 ± 0.01

观测值在一个标准误差范围内与理论预期相符。

例：含参数的矩方法估计

假设理论预言 $e^+e^- \rightarrow \mu^+\mu^-$ 的角分布含未知参数 α ：

$$\frac{dn}{d\cos\theta} \propto n(1 + \alpha \cos^2 \theta)$$

计算 $\cos\theta$ 二阶原点矩的数学期望值

$$E[\cos^2 \theta] = \frac{\int_{-1}^{+1} \cos^2 \theta \cdot n(1 + \alpha \cos^2 \theta) d\cos \theta}{\int_{-1}^{+1} n(1 + \alpha \cos^2 \theta) d\cos \theta} = \frac{5 + 3\alpha}{5(3 + \alpha)}$$

则参数 α 与二阶原点矩的关系为

$$\alpha = \frac{5(3E[\cos^2 \theta] - 1)}{3 - 5E[\cos^2 \theta]}$$



$$\hat{\alpha} = \frac{5(3\overline{\cos^2 \theta} - 1)}{3 - 5\overline{\cos^2 \theta}}$$

只要函数是可积的，采用矩方法原则上就可以确定参数。

矩方法应用的问题

非物理解问题：利用矩方法测定参数，可能会出现非物理结果。例如前例的二阶代数矩中，如果

$$\overline{\cos^2 \theta} \rightarrow 0.6 \quad \rightarrow \quad \alpha \rightarrow \infty$$

在矩方法中，我们无法加上限制条件使得参数的估计值保持在物理允许的范围。

假设检验问题：利用矩方法测定参数，由于只比较积分值并解方程得到参数估计值，信息含量不足，因此无法判断得到的参数是否合理。实际应用中需要辅之以其它方法来检验。

适用范围问题：矩方法虽然简单，但在处理多参数问题中，由于涉及更高阶的积分，使研究变得复杂。在这种情况下，可以考虑采用所谓的“广义矩方法”。

最大似然法、最小二乘法和矩方法

	矩方法	最大似然法	最小二乘法
数据输入	单个事例	单个事例	直方图
多维问题	最容易	归一化较复杂	较难
充分性	会有信息丢失	最具充分性	有时与区间大小有关
相合性	收敛于真值	收敛于真值	收敛于真值
有效性	不是最有效	通常最有效	基本上与似然法一样
无偏性	渐进无偏	渐进无偏	渐进无偏
拟合优度	较难评估	较难评估	很容易

充分性：估计量应包含观测值对于未知参数的全部信息；

相合性：样本容量增大时，估计值收敛于真值；

有效性：估计量的分布对其期望值具有最小方差；

无偏性：无论样本容量多大，估计值与真值无系统偏差。

小结

- 矩的定义
- 矩方法估计量及其应用
- 矩方法与最大似然法和最小二乘法的比较