

曲线拟合

张松

核科学与技术系|现代物理研究所





提纲

- 最小二乘法原理
- 线性参数的最小二乘拟合



第一节 引言

- 实验目的：测量数据，基于测量数据的物理模型
- 拟合目的
 - ✓ 物理解释
 - ✓ 理论参数确定
- 方法



引言 (续)

- 已知某未知量 x 和 y 之间存在函数关系,
 $y = f(x; c_1, c_2, \dots, c_m)$, 为确定参数 c_1, c_2, \dots, c_m 的最佳值或说寻求上述理论曲线的最佳估计值
- 实验物理中, 物理量 x 和 y 之间关系未知, 需要得到其间的经验公式, 通过对分布 (测量结果) 的观察, 可以用诸如多项式、三角函数、e指数等形式进行拟合
- 参数估计实际上也是一种拟合



引言 (续)

- 拟合标准：确定明确的标准，使得拟合得到的参数 $\bar{c}_1, \bar{c}_2, \dots, \bar{c}_m$ 具有某种最佳性质。这种性质一般可以表示为使得某目标函数 $d = d(x_1, x_2, \dots, x_N, y_1, y_2, \dots, y_N; c_1, c_2, \dots, c_m)$ 取极值。
- 注意：实际问题中，以可以忽略测量误差或精度更高的变量作为自变量 x ，则观测量 x^* 可以视为 x 的准确值 x_i 。对应于某个固定的 x_i 值， y 的观测值 y_i^* 是一个随机变量。如果测量的系统误差可以忽略， y_i^* 的期待值等于其理论值，
$$\langle y_i^* \rangle = y_i = f(x_i; c_1, \dots, c_m)$$



第二节 最小二乘原理

- 最小二乘准则
- 最小二乘法与最大似然法



最小二乘准则

- 使用最广泛的曲线拟合方法
- **目标函数**：各观测点残差的加权平方和作为目标函数，观测点 i 的**残差** δ_i 定义为 y 的观测值与理论值（拟合值）之差，
$$\delta_i \equiv y_i^* - y_i = y_i^* - f(x_i; c_1, \dots, c_m)$$

若观测值 y_i^* 的标准误差为 σ_i ，则观测值 y_i^* 的**权重因子**为 $\omega_i = 1/\sigma_i^2$

拟合的 χ^2 量定义为：观测点残差的加权平方和，

$$\chi^2 \equiv \sum_{i=1}^N \omega_i \delta_i^2 = \sum_{i=1}^N \omega_i \left[y_i^* - f(x_i; \underline{c}) \right]^2$$

- 最小二乘法就是用拟合的 χ^2 量作为目标函数，寻求使得 χ^2 量最小的参数值 $\hat{\underline{c}}$ 作为参数的估计值，
$$\chi^2|_{\underline{c}=\hat{\underline{c}}} = \sum_{i=1}^N \omega_i \delta_i^2|_{\underline{c}=\hat{\underline{c}}} = \min$$



最小二乘准则 (续)

- χ^2 量的极值可由函数极值法给出, 来确定参数值 $\underline{\hat{c}}$,

$$\frac{\partial \chi^2}{\partial c_k} = \frac{\partial}{\partial c_k} \sum_{i=1}^N \omega_i \delta_i^2 = \sum_{i=1}^N \omega_i 2\delta_i \frac{\partial \delta_i}{\partial c_k} = 0 \quad (k = 1, \dots, m)$$

- 将 δ_i 定义带入上式,

$$\sum_{i=1}^N \omega_i \left[y_i^* - f(x_i; c_1, \dots, c_m) \right] \frac{\partial f(x_i; c_1, \dots, c_m)}{\partial c_k} = 0 \quad (k = 1, \dots, m)$$

, 可解出参数的最小二乘估计值 $\hat{c}_1, \dots, \hat{c}_m$

- 上式, 由 m 个方程组成, 叫正则方程。当理论曲线 $y = f(x; \underline{c})$ 是对于参数 \underline{c} 的线性函数时, 上述方程组是未知参数 \underline{c} 的线性联立方程组, 可直接解出



最小二乘法与最大似然法

- 由中心极限定理可知，对应于每一个 x_i ， y 的观测值 y_i^* 服从正态分布，分布的期待值应为理论值 $f(x_i; \underline{c})$ ，方差为 σ_i^2 ，则观测值 $(y_1^*, y_2^*, \dots, y_N^*)$ 的似然函数为（各次测量独立进行）

$$\begin{aligned} L = p(\underline{y}^*; \underline{c}) &= \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left\{ -\frac{1}{2} \left[\frac{y_i^* - f(x_i; \underline{c})}{\sigma_i} \right]^2 \right\} \\ &= \frac{1}{(2\pi)^{N/2} \sigma_1 \cdots \sigma_N} \exp \left\{ -\frac{1}{2} \sum_{i=1}^N \frac{[y_i^* - f(x_i; \underline{c})]^2}{\sigma_i^2} \right\} \end{aligned}$$

- 最大似然法的目标函数为似然函数，参数 \underline{c} 的最大似然估计值 $\hat{\underline{c}}$ 使似然函数取最大值，

$$L|_{\underline{c}=\hat{\underline{c}}} = \max, \text{ 等价于 } \sum_{i=1}^N \frac{1}{\sigma_i^2} [y_i^* - f(x_i; \underline{c})]^2 \Big|_{\underline{c}=\hat{\underline{c}}} = \min$$

- 当观测值服从正态分布时，最小二乘法同最大似然法一致，最小二乘估计值具有最大似然估计值的最优性质。当观测值不服从正态分布时，最小二乘估计并不一定是参数的最大似然估计



最小二乘法与最大似然法 (续)

- 对于正态变量 $y_i^* (1, 2, \dots, N)$,

$$\chi^2 = \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[y_i^* - f(x_i; \underline{c}) \right]^2, \text{ 服从自由度为 } N \text{ 的 } \chi^2 \text{ 分布。}$$

把 χ^2 量中的 m 个未知参数用它们的最小二乘估计值 $\hat{\underline{c}}$ 代入, 可

$$\text{得最小 } \chi^2 \text{ 值 } \chi_{min}^2 = \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[y_i^* - f(x_i; \hat{c}_1, \dots, \hat{c}_m) \right]^2$$

- 显然, 最小 χ^2 值服从自由度为 $N - m$ 的 χ^2 分布,
 $\chi_{min}^2 \sim \chi^2(N - m)$, χ_{min}^2 期待值应该是 $N - m$, 可用来检验
曲线拟合的质量, $N - m$ 叫做拟合的自由度, χ_{min}^2 称为拟合
的优度因子



第三节 线性参数的最小二乘拟合

线性情况理论曲线：

$y = f(x; \underline{c}) = y_0(x) + c_1 f_1(x) + c_2 f_2(x) + \cdots + c_m f_m(x)$, 其中 $y_0(x)$ 和 $f_j(x) (j = 1, 2, \dots, m)$ 是 x 的已知函数

- 参数的估计值
- 参数估计的误差
- 观测值的重新估计
- 拟合曲线的误差
- 等精度测量时观测值的方差估计
- 非独立观测值的最小二乘拟合



参数的估计

- 对于 N 个观测点 $(x_1, y_1^*), (x_2, y_2^*), \dots, (x_N, y_N^*)$, 参数 \underline{c} 的最小二乘拟合值可由第一节的方程组解出,

$$\sum_{i=1}^N \omega_i \left[y_i^* - f(x_i; c_1, \dots, c_m) \right] \frac{\partial f(x_i; c_1, \dots, c_m)}{\partial c_k} = 0 \quad (k = 1, \dots, m)$$

, 将线性函数 $y = f(x; c) = y_0(x) + c_1 f_1(x) + c_2 f_2(x) + \dots + c_m f_m(x)$ 代入, 即得到 \underline{c} 的线性方程组,

$$\sum_{i=1}^N \omega_i \left\{ y_i^* - \left[y_0(x_i) + \sum_{j=1}^m c_j f_j(x_i) \right] \right\} f_k(x_i) = 0 \quad (k = 1, 2, \dots, m)$$

- 展开、移项、整理:

$$\sum_{j=1}^m c_j \sum_{i=1}^N \omega_i f_j(x_i) f_k(x_i) = \sum_{i=1}^N \omega_i \left[y_i^* - y_0(x_i) \right] f_k(x_i) \quad (k = 1, 2, \dots, m)$$



参数的估计 (续)

考察左边,

$$\rightarrow \begin{pmatrix} c_1 \sum_{i=1}^N \omega_i f_1(x_i) f_1(x_i) + c_2 \sum_{i=1}^N \omega_i f_2(x_i) f_1(x_i) + \cdots + c_m \sum_{i=1}^N \omega_i f_m(x_i) f_1(x_i) \\ c_1 \sum_{i=1}^N \omega_i f_1(x_i) f_2(x_i) + c_2 \sum_{i=1}^N \omega_i f_2(x_i) f_2(x_i) + \cdots + c_m \sum_{i=1}^N \omega_i f_m(x_i) f_2(x_i) \\ \cdots \\ c_1 \sum_{i=1}^N \omega_i f_1(x_i) f_m(x_i) + c_2 \sum_{i=1}^N \omega_i f_2(x_i) f_m(x_i) + \cdots + c_m \sum_{i=1}^N \omega_i f_m(x_i) f_m(x_i) \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} \sum_{i=1}^N \omega_i f_1(x_i) f_1(x_i) & \sum_{i=1}^N \omega_i f_2(x_i) f_1(x_i) & \cdots & \sum_{i=1}^N \omega_i f_m(x_i) f_1(x_i) \\ \sum_{i=1}^N \omega_i f_1(x_i) f_2(x_i) & \sum_{i=1}^N \omega_i f_2(x_i) f_2(x_i) & \cdots & \sum_{i=1}^N \omega_i f_m(x_i) f_2(x_i) \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{i=1}^N \omega_i f_1(x_i) f_m(x_i) & \sum_{i=1}^N \omega_i f_2(x_i) f_m(x_i) & \cdots & \sum_{i=1}^N \omega_i f_m(x_i) f_m(x_i) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}$$



参数的估计 (续)

$$\rightarrow \begin{pmatrix} \omega_1 f_1(x_1) & \omega_2 f_1(x_2) & \cdots & \omega_N f_1(x_N) \\ \omega_1 f_2(x_1) & \omega_2 f_2(x_2) & \cdots & \omega_N f_2(x_N) \\ \cdots & & & \\ \omega_1 f_m(x_1) & \omega_2 f_m(x_2) & \cdots & \omega_N f_m(x_N) \end{pmatrix} \begin{pmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_m(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_m(x_2) \\ \cdots & & & \\ f_1(x_N) & f_2(x_N) & \cdots & f_m(x_N) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} f_1(x_1) & f_1(x_2) & \cdots & f_1(x_N) \\ f_2(x_1) & f_2(x_2) & \cdots & f_2(x_N) \\ \cdots & \cdots & \cdots & \cdots \\ f_m(x_1) & f_m(x_2) & \cdots & f_m(x_N) \end{pmatrix} \begin{pmatrix} \omega_1 & 0 & \cdots & 0 \\ 0 & \omega_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \omega_N \end{pmatrix} \begin{pmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_m(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_m(x_2) \\ \cdots & & & \\ f_1(x_N) & f_2(x_N) & \cdots & f_m(x_N) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}$$



参数的估计 (续)

定义矩阵,

$$\underline{C} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}, \quad \underline{Y}^* = \begin{pmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_N^* \end{pmatrix}, \quad \underline{Y}_0 = \begin{pmatrix} y_0(x_1) \\ y_0(x_2) \\ \vdots \\ y_0(x_N) \end{pmatrix},$$

$$\underline{F} = \begin{pmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_m(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_m(x_2) \\ \cdots & \cdots & \cdots & \cdots \\ f_1(x_N) & f_2(x_N) & \cdots & f_m(x_N) \end{pmatrix} = (f_{ij}), \quad f_{ij} = f_j(x_i)$$

$$\underline{W}_Y = \underline{V}_Y^{-1} = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sigma_m^2 \end{pmatrix}^{-1} = \begin{pmatrix} \omega_1 & 0 & \cdots & 0 \\ 0 & \omega_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \omega_N \end{pmatrix}$$

\underline{W}_Y 称为权重矩阵, \underline{V}_Y 是观测值 y_i^* 的协方差阵



参数的估计 (续)

对 $\sum_{j=1}^m c_j \sum_{i=1}^N \omega_i f_j(x_i) f_k(x_i) = \sum_{i=1}^N \omega_i [y_i^* - y_0(x_i)] f_k(x_i) \quad (k = 1, 2, \dots, m)$ 右侧做类似操作，则可以用矩阵形式写成， $(\underline{F}^T \underline{W}_Y \underline{F}) \underline{C} = \underline{F}^T \underline{W}_Y (\underline{Y}^* - \underline{Y}_0)$

解上述方程组或矩阵方程，可得参数 \underline{C} 的最小二乘估计值，

$$\underline{\hat{C}} = \begin{pmatrix} \hat{c}_1 \\ \hat{c}_2 \\ \vdots \\ \hat{c}_m \end{pmatrix} = (\underline{F}^T \underline{W}_Y \underline{F})^{-1} \underline{F}^T \underline{W}_Y (\underline{Y}^* - \underline{Y}_0)$$

可求得最小二乘估计值 $\underline{\hat{C}}$ 的期待值，

$$\begin{aligned} \langle \underline{\hat{C}} \rangle &= (\underline{F}^T \underline{W}_Y \underline{F})^{-1} \underline{F}^T \underline{W}_Y \langle (\underline{Y}^* - \underline{Y}_0) \rangle \\ &= (\underline{F}^T \underline{W}_Y \underline{F})^{-1} (\underline{F}^T \underline{W}_Y \underline{F}) \langle \underline{C} \rangle = \underline{C} \end{aligned}$$

因此参数的最小二乘估计是无偏估计



参数估计的误差

$$\underline{\hat{C}} = \begin{pmatrix} \hat{c}_1 \\ \hat{c}_2 \\ \vdots \\ \hat{c}_m \end{pmatrix} = \left(\underline{F}^T \underline{W}_Y \underline{F} \right)^{-1} \underline{F}^T \underline{W}_Y \left(\underline{Y}^* - \underline{Y}_0 \right), \text{ 在线性情况下, 参数 } \underline{C} \text{ 的最小二乘}$$

估计值 $\underline{\hat{C}}$ 是观测值 \underline{Y}^* 的线性函数, 利用线性函数的误差传播公式, 可以由观测值 \underline{Y}^* 的协方差阵 \underline{V}_Y 计算 **参数最小二乘估计值 $\underline{\hat{C}}$ 的协方差阵**,

$$\underline{V}_{\hat{C}} = \left[\left(\underline{F}^T \underline{W}_Y \underline{F} \right)^{-1} \underline{F}^T \underline{W}_Y \right] \underline{V}_Y \left[\left(\underline{F}^T \underline{W}_Y \underline{F} \right)^{-1} \underline{F}^T \underline{W}_Y \right]^T, \text{ 因 } \underline{W}_Y \text{ 是对称矩阵,}$$

$\underline{W}_Y^T = \underline{W}_Y$, 故 $\left(\underline{F}^T \underline{W}_Y \underline{F} \right)$ 也是对称矩阵, 利用矩阵运算的性质:

$$\left(\left(\underline{A} \underline{B} \cdots \underline{Z} \right)^T = \underline{Z}^T \cdots \underline{B}^T \underline{A}^T \quad \left(\underline{A}^{-1} \right)^T = \left(\underline{A}^T \right)^{-1} \quad \left(\underline{A} \underline{B} \cdots \underline{Z} \right)^{-1} = \underline{Z}^{-1} \cdots \underline{B}^{-1} \underline{A}^{-1} \right)$$

, 求 $\underline{\hat{C}}$ 的协方差阵公式简化为: $\underline{V}_{\hat{C}} = \left(\underline{F}^T \underline{W}_Y \underline{F} \right)^{-1}$, **协方差阵 $\underline{V}_{\hat{C}}$ 的第 j 个对角元素**

的平方根就是参数估计值 \hat{c}_j 的标准误差。一般地说, $\underline{V}_{\hat{C}}$ 并非对角矩阵, 各参数估计值彼此相关, $\underline{V}_{\hat{C}}$ 的第 ij 个矩阵元就是 \hat{c}_i 和 \hat{c}_j 的协方差, 比较完全的表达参数估计值误差的办法, 是列出协方差阵 $\underline{V}_{\hat{C}}$



观测值的重新估计

- 把参数估计值 $\hat{\underline{C}}$ 代入理论曲线
 $y = f(x; \underline{c}) = y_0(x) + c_1 f_1(x) + c_2 f_2(x) + \cdots + c_m f_m(x)$, 得到理论曲线的最小二乘估计, $y = f(x; \hat{\underline{C}}) = y_0(x) + \hat{c}_1 f_1(x) + \hat{c}_2 f_2(x) + \cdots + \hat{c}_m f_m(x)$
- 用 \hat{y}_i 表示在拟合曲线上对应于 x_i 的y坐标值,
 $\hat{y}_i = f(x_i; \hat{\underline{C}}) = y_0(x_i) + \hat{c}_1 f_1(x_i) + \hat{c}_2 f_2(x_i) + \cdots + \hat{c}_m f_m(x_i) \quad (i = 1, 2, \dots, N)$,
或写成矩阵形式, $\hat{\underline{Y}} = \underline{Y}_0 + \underline{F}\hat{\underline{C}}$, 其中 $\hat{\underline{Y}} = \begin{pmatrix} \hat{y}_1 & \hat{y}_2 & \cdots & \hat{y}_N \end{pmatrix}^T$ 。表明 $\hat{\underline{Y}}$ 是参数估计值 $\hat{\underline{C}}$ 的线性函数, 由线性函数的误差传播公式, 估计值 $\hat{\underline{Y}}$ 的协方差为, $\underline{V}_{\hat{\underline{Y}}} = \underline{F}\underline{V}_{\hat{\underline{C}}}\underline{F}^T$ 。
计算出的协方差阵 $\underline{V}_{\hat{\underline{Y}}}$ 对角线上的第*i*个元素就是对应的估计值 \hat{y}_i 的方差 $\sigma^2(\hat{y}_i)$
- 拟合的 χ^2 量也可以写成矩阵的形式,
$$\chi^2 = \left(\underline{Y}^* - \underline{Y}_0 - \underline{F}\underline{C} \right)^T \underline{W}^T \left(\underline{Y}^* - \underline{Y}_0 - \underline{F}\underline{C} \right)$$

拟合的最小值为,

$$\chi^2_{\min} = \left(\underline{Y}^* - \underline{Y}_0 - \underline{F}\underline{C} \right)^T \underline{W}^T \left(\underline{Y}^* - \underline{Y}_0 - \underline{F}\underline{C} \right) = \sum_{i=1}^N \omega_i \left[y_i^* - f(x_i; \hat{\underline{C}}) \right]^2$$



拟合曲线的误差

- 拟合曲线上对应任意 x 的函数值 y 由

$y = f(x; \underline{\hat{C}}) = y_0(x) + \hat{c}_1 f_1(x) + \hat{c}_2 f_2(x) + \cdots + \hat{c}_m f_m(x)$ 式算出,
其方差可采用误差传递公式计算,

$$\sigma^2(y) \approx \sum_{i=1}^N \left(\frac{\partial f}{\partial x_i} \right)_{\underline{x}=\langle \underline{x} \rangle}^2 \sigma^2(x_i) + 2 \sum_{i < j} \left(\frac{\partial y}{\partial x_i} \frac{\partial y}{\partial x_j} \right)_{\underline{x}=\langle \underline{x} \rangle} \text{Cov}(x_i, x_j)$$

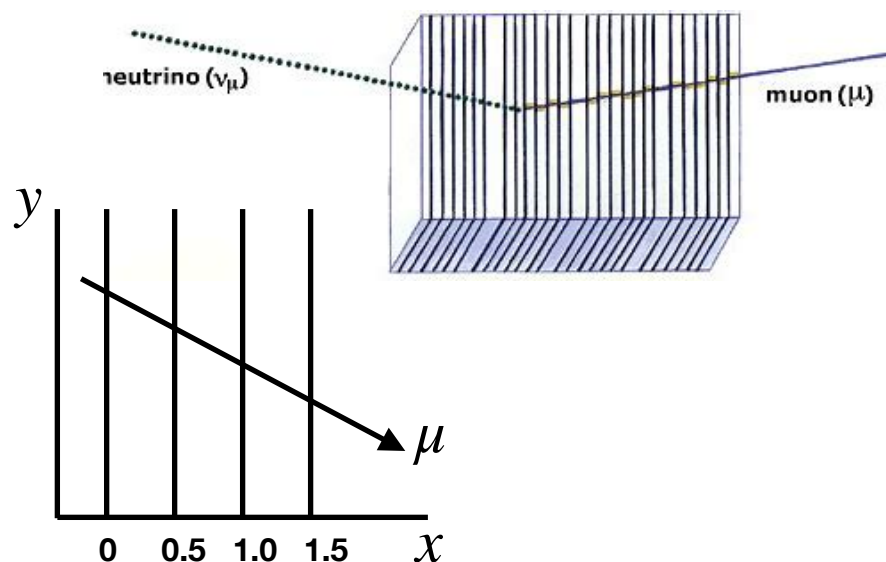
- 这里, y 应是 $f(x; \underline{\hat{C}})$, x_i 就是 c_i , 拟合曲线上对应任意 x 的函数值 y 的方差应是,

$$\sigma^2(y) \approx \sum_{i=1}^m \left(\frac{\partial f}{\partial c_i} \right)_{\underline{c}=\underline{\hat{c}}}^2 \sigma^2(c_i) + 2 \sum_{i < j} \left(\frac{\partial y}{\partial c_i} \frac{\partial y}{\partial c_j} \right)_{\underline{c}=\underline{\hat{c}}} \text{Cov}(c_i, c_j)$$

- 也可以利用 $\underline{V}_{\hat{Y}} = \underline{F} \underline{V}_{\hat{C}} \underline{F}^T$ 进行计算

例题 粒子径迹的拟合

用多板火花室测定高能 μ 子径迹， μ 子的真实径迹是直线。如图所示， x 代表垂直于极板方向的坐标， y 代表沿极板方向的坐标，则在 xy 平面内粒子径迹的理论曲线是直线
 $y = c_1 + c_2x$ ，设第 i 块阴极板的坐标为 x_i ，不考虑其误差。在第 i 块阴极板处 μ 子径迹位置的观测值为 y_i^* ，其标准误差为 $\sigma_i(y_i^*)$ ，火花室总共由4块阴极板构成， x_i 、 y_i^* 和 $\sigma_i(y_i^*)$ 的数据见下表，试做数据分析。



i	1	2	3	4
x_i	0	0.5	1.0	1.5
y_i^*	1.61	1.32	0.80	0.70
$\sigma(y_i^*)$	0.14	0.14	0.14	0.14



例题 粒子径迹的拟合 (续)

该理论曲线 $y = c_1 + c_2x$ 是参数 c_1 和 c_2 的线性函数, 其中 $y_0(x) = 0$, $f_1(x) = 1$, $f_2(x) = x$, 依次定义各矩阵,

$$\underline{C} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \quad \underline{Y}^* = \begin{pmatrix} 1.61 \\ 1.32 \\ 0.80 \\ 0.70 \end{pmatrix}, \quad \underline{Y}_0 = 0, \quad \underline{F} = \begin{pmatrix} 1 & 0 \\ 1 & 0.5 \\ 1 & 1 \\ 1 & 1.5 \end{pmatrix}$$

$$\underline{V}_Y = \begin{pmatrix} 0.14^2 & 0 & 0 & 0 \\ 0 & 0.14^2 & 0 & 0 \\ 0 & 0 & 0.14^2 & 0 \\ 0 & 0 & 0 & 0.14^2 \end{pmatrix} = 2.0 \times 10^{-2}, \quad \underline{W}_Y = \underline{V}_Y^{-1} = 50I$$

上式中 I 为 4×4 的单位矩阵, 根据参数计算公式计算 \underline{C} 的最小二乘值

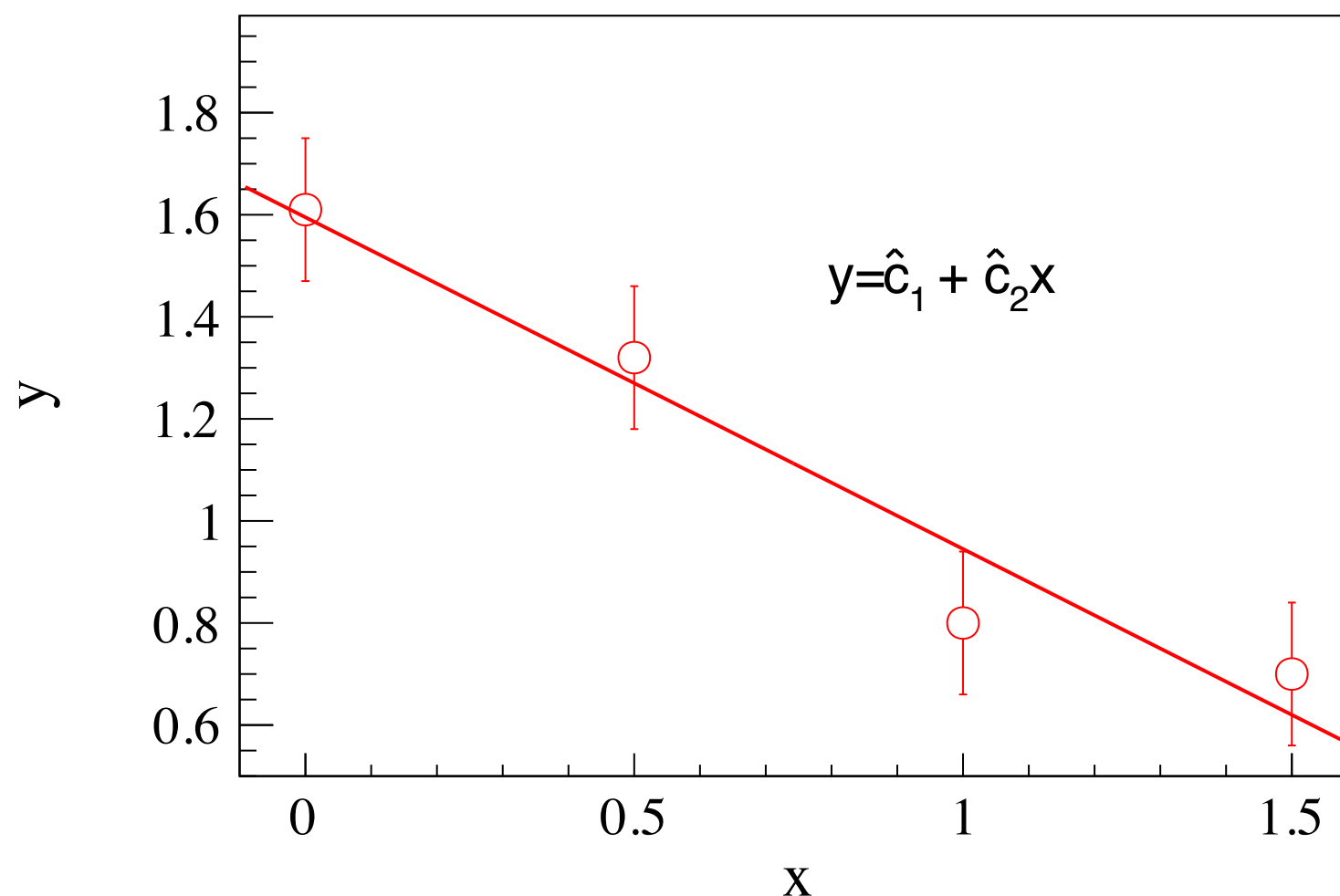
$$\begin{aligned} \underline{\hat{C}} &= \begin{pmatrix} \hat{c}_1 \\ \hat{c}_2 \end{pmatrix} = \left(\underline{F}^T \underline{W}_Y \underline{F} \right)^{-1} \underline{F}^T \underline{W}_Y \left(\underline{Y}^* - \underline{Y}_0 \right) \\ &= \left(\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0.5 & 1 & 1.5 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0.5 \\ 1 & 1 \\ 1 & 1.5 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0.5 & 1 & 1.5 \end{pmatrix} \begin{pmatrix} 1.61 \\ 1.32 \\ 0.80 \\ 0.70 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 3 \\ 3 & 3.5 \end{pmatrix}^{-1} \begin{pmatrix} 4.43 \\ 2.51 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 3.5 & -3 \\ -3 & 4 \end{pmatrix} \begin{pmatrix} 4.43 \\ 2.51 \end{pmatrix} = \begin{pmatrix} 1.60 \\ -0.65 \end{pmatrix} \end{aligned}$$



例题 粒子径迹的拟合（续）

参数 c_1 和 c_2 的估计值为 $\hat{c}_1 = 1.60$, $\hat{c}_2 = -0.65$

把 \hat{c}_1 、 \hat{c}_2 带入 $y = c_1 + c_2x$ 得到 μ 子的径迹的拟合结果为：
 $y = 1.60 - 0.65x$ ，下图为观测数据和拟合结果。





等精度测量时观测值的方差估计

- 以上讨论为非等精度测量的一般情况
- 对于等精度测量，各观测值的权重 ω_i 都相等，参数 \underline{C} 的最小二乘估计式简化为

$$\hat{\underline{C}} = (\underline{F}^T \underline{F})^{-1} \underline{F}^T (\underline{Y}^* - \underline{Y}_0)$$

这样可以直接算出参数 \underline{C} 的估计值 $\hat{\underline{C}}$ 。把 $\hat{\underline{C}}$ 代入 $f(x; \underline{C})$ 中可求得各观测点的重新估计值 $\hat{\underline{Y}}$ ，进而计算出各观测值的残差平方和 R ：

$$R = \sum_{i=1}^N \delta_i^2 = \sum_{i=1}^N [y_i^* - f(x; \hat{\underline{C}})]^2$$

最小 χ^2 值为

$$\chi_{min}^2 = \frac{R}{\sigma_y^2} \sim \chi^2(N - m)$$

故有

$$\left\langle \frac{R}{\sigma_y^2} \right\rangle = N - m$$

上式中， σ_y^2 是各观测值 y_i^* 的方差。因此，在等精度测量时，如果观测值 \underline{Y}^* 的方差 σ_y^2 未知，可由下式估计得到：

$$\hat{\sigma}_y^2 = \frac{R}{N - m} = \frac{\sum_{i=1}^N [y_i^* - f(x; \hat{\underline{C}})]^2}{N - m}$$

将上式代入 \underline{W}_y 矩阵，可得到协方差阵 $\underline{V}_{\hat{\underline{C}}}$ ，进而可得 $\hat{\underline{Y}}$ 的协方差阵 $\underline{V}_{\hat{\underline{Y}}}$ ，其对角元素即 $\hat{\underline{Y}}$ 的方差 $\sigma^2(\hat{y})$ 。由于 $\sigma^2(\hat{y})$ 纯系样本估计得到，若较严格考虑，当 $N - m$ 较小(≤ 10)时，结果应按 t 分布报道， t 分布的自由度为 $N - m$ 。



非独立观测值的最小二乘拟合

- 前面所讨论的是独立观测的情况，各个观测值彼此无关，协方差阵 \underline{V}_y 和权重 \underline{W}_y 都是对角矩阵。
- 对于非独立观测的情况，各观测值之间的协方差不等于零，因此协方差阵 \underline{V}_y 和权重 \underline{W}_y 都不是对角矩阵。但是，对于独立观测值的线性最小二乘拟合的公式可以全部推广到对于非独立观测值的拟合。



课后习题

- 计算例题中，参数估计值 \hat{c}_1 、 \hat{c}_2 的标准误差
- 计算例题中，各极板处径迹位置的估计值及其标准误差



ROOT拟合初步

- TF1类对TH1的拟合，比较简单
- 多条曲线（幅度不同），但具有其它相同的参数，TMinuit

ROOT拟合初步 (续)

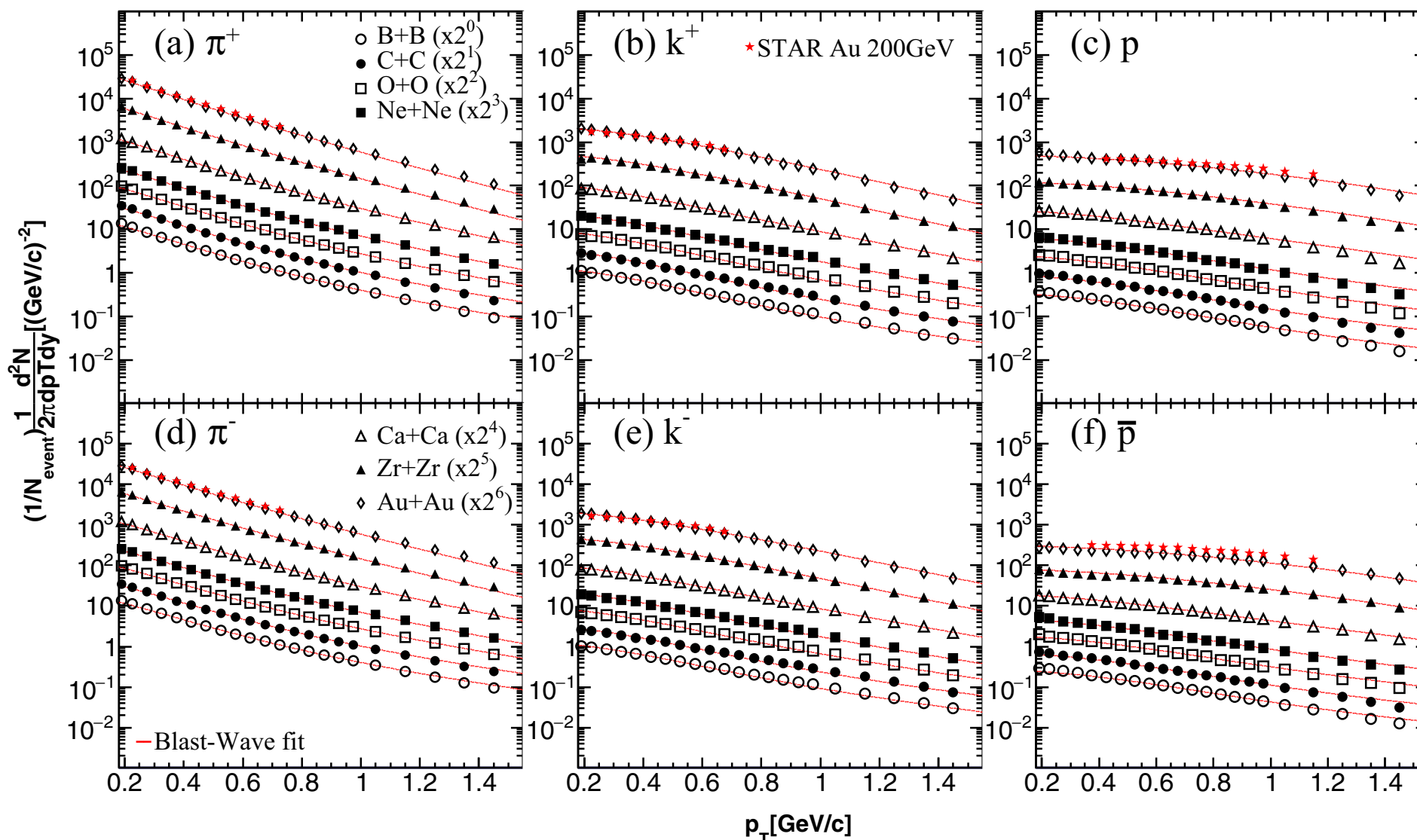


FIG. 1. Transverse momentum p_T spectra at midrapidity ($|y| < 0.1$) for π^\pm , k^\pm , p , and \bar{p} in $^{10}\text{B} + ^{10}\text{B}$, $^{12}\text{C} + ^{12}\text{C}$, $^{16}\text{O} + ^{16}\text{O}$, $^{20}\text{Ne} + ^{20}\text{Ne}$, $^{40}\text{Ca} + ^{40}\text{Ca}$, $^{96}\text{Zr} + ^{96}\text{Zr}$, and $^{197}\text{Au} + ^{197}\text{Au}$ collision systems at $\sqrt{s_{NN}} = 200$ GeV. The spectra are scaled by a factor of 2^n , as marked in the inset (note that the data of the boron collision system are not scaled). The curves indicate that the blast-wave model combination fits the 0%–5% centrality AMPT results for different collision systems. Experimental data are taken from the STAR Collaboration for Au + Au collisions at $\sqrt{s_{NN}} = 200$ GeV [14].

$$\frac{1}{p_T} \frac{dN}{dp_T} \propto \int_0^R r dr m_T I_0 \left(\frac{p_T \sinh \rho}{T_{\text{kin}}} \right) K_1 \left(\frac{m_T \cosh \rho}{T_{\text{kin}}} \right)$$