



# 实验物理中的统计方法

## 第十一章：解谱法

杨振伟

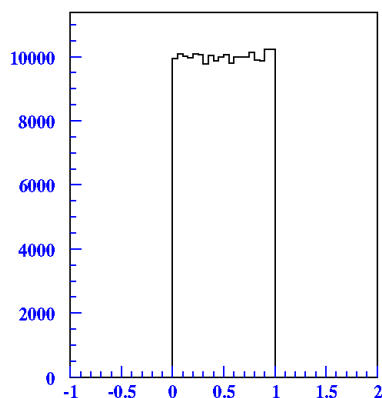
# 本章要点

---

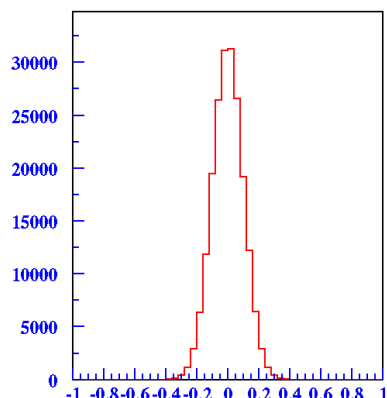
- 数学描述, 响应矩阵 (函数)
- 求响应矩阵的逆
- 修正因子
- 正规化的解谱法
  - a) Tikhonov 规则
  - b) MaxEnt 规则
- 估计量的方差与偏倚
- 规范化参数的选择
- 举例

# 图像还原问题

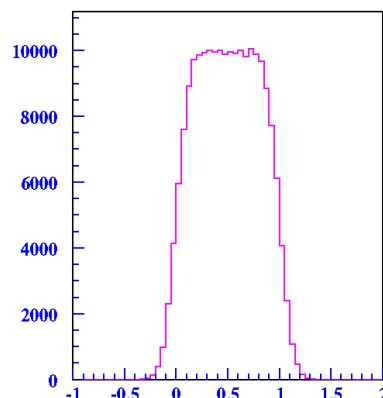
一个常见的问题：由于实验仪器的原因导致图像变形，例如



真实分布



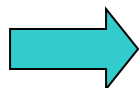
探测器响应



实验观测分布

如果已知探测器响应（可通过探测器模拟得到其形式），

能否还原出不受实验仪器影响的真实分布？



Unfolding(解谱法)

# 解谱问题的表述(1)

考虑随机变量 $y$ ，目标：寻找概率密度函数 $f(y)$

1) 如果已知参数化形式  $f(y; \vec{\theta})$

最大似然法  $\Rightarrow \hat{\vec{\theta}} \Rightarrow f(y; \hat{\vec{\theta}})$

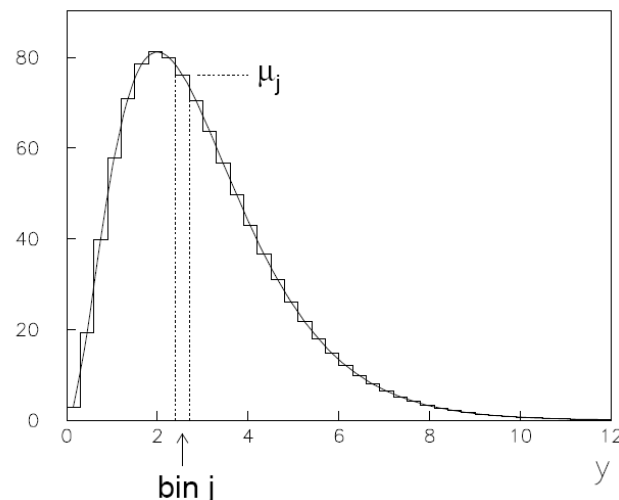
2) 不知概率密度的形式，构造直方图

直方图分  $M$  个区间，概率密度离散化  
第 $j$ 个区间的概率为

$$p_j = \int_{\text{bin } j} f(y) dy, \quad j = 1, \dots, M$$

$$\mu_j = \mu_{\text{tot}} p_j \quad \leftarrow \text{“真实的直方图”}$$

为 $\mu_j$ (或 $p_j$ )构造估计量  $\Rightarrow$  离散化的p.d.f.  
(参数的数目 = 区间的数目  $M$ )



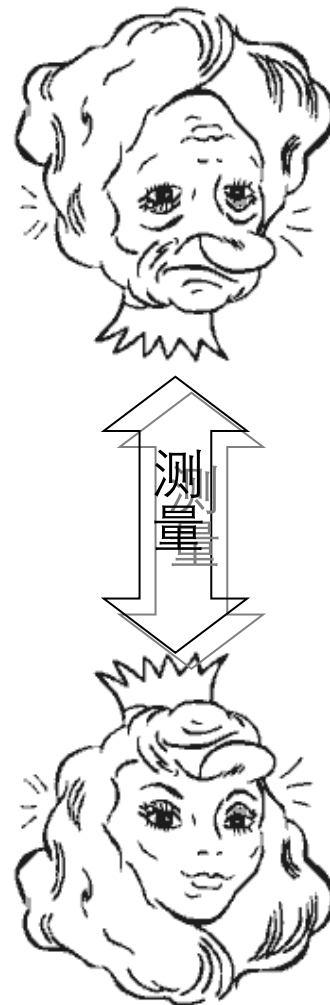
# 解谱问题的表述(2)

问题：  $y$  的测量不可能没有误差

- ➡
- 1) 第 $i$ 个区间的 $y$ 在第 $j$ 个区间测量到  
(测量分辨率)
  - 2) 第 $i$ 个区间的某些事例没有测量到  
(测量效率)
  - 3) 某些区间的事例完全测量不到  
(接受度)

后果：  $f(y)$  模糊化，峰被展宽，  
甚至分布形状完全变形。

后果严重时需要考虑解谱法还原真实分布



# 响应矩阵

$y$ : 真值;  $x$ : 观测值      测量的影响可用积分方程表示:

$$f_{\text{meas}}(x) = \int R(x|y) f_{\text{true}}(y) dy$$

$R(x|y)$ : 响应函数, 表示真值为 $y$ 观测值为 $x$ 的概率

$R(x|y)$  对 $y$ 求积分  $\rightarrow$  测量值为 $x$ 的概率

离散化

真实直方图

观测直方图  
(期待值)

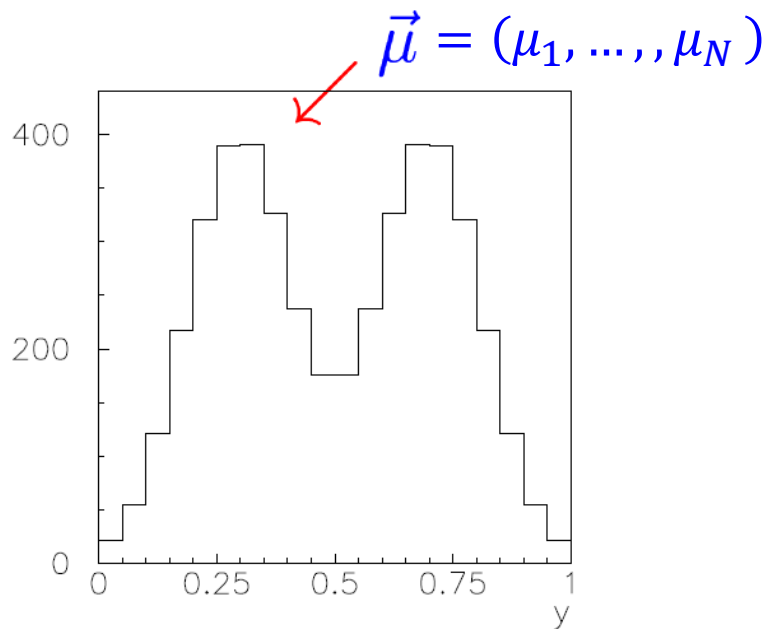
$$v_i = \sum_{j=1}^M R_{ij} \mu_j, \quad i = 1, \dots, N$$

响应矩阵  $R_{ij} = P(\text{观测值在第}i\text{区}|\text{真实值在第}j\text{区})$

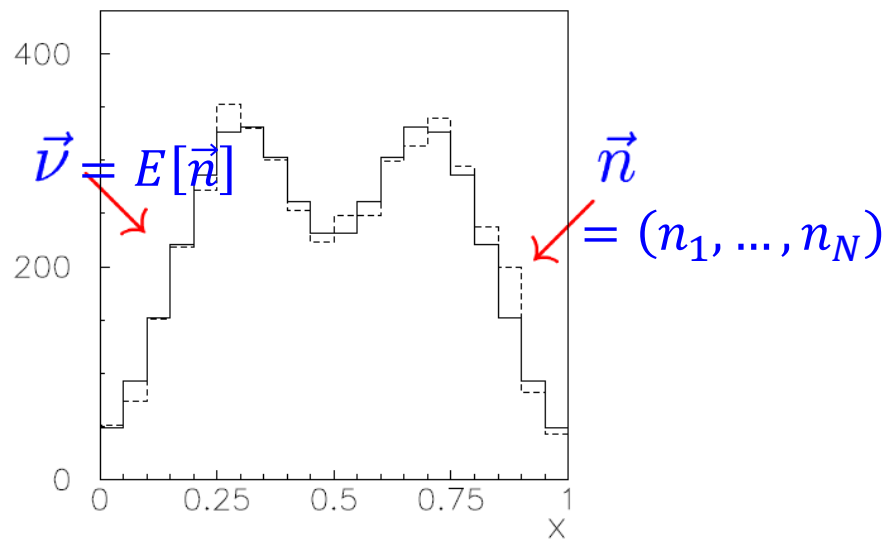
# 响应矩阵

响应矩阵  $R_{ij} = P(\text{观测值在第}i\text{区}|\text{真实值在第}j\text{区})$

真实直方图:  $\vec{\mu}$   
离散化的p.d.f.



观测数据  $\vec{n} = (n_1, \dots, n_N)$   
和数据的期待值  $\vec{v} = E[\vec{n}]$



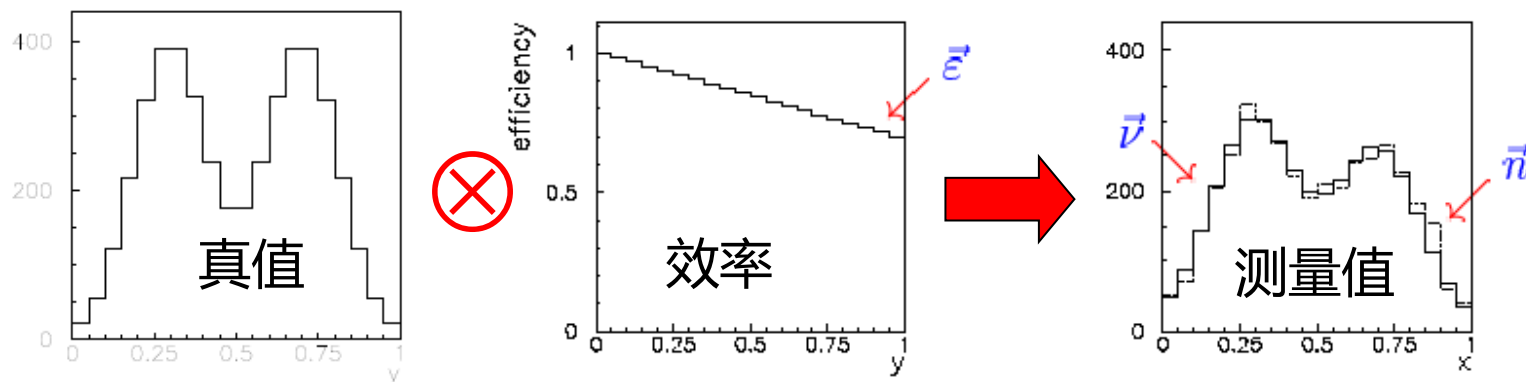
注意:  $\vec{\mu}, \vec{v}$  是常数, 而  $\vec{n}$  会受到统计涨落的影响。

# 效率、本底

有时，事例可能没有被探测到: 效率

$$\sum_{j=1}^M R_{ij} = \sum_{i=1}^N P(\text{观测值在第} i \text{区} | \text{真实值在第} j \text{区})$$
$$= P(\text{观测值在全范围} | \text{真实值在第} j \text{区}) = \varepsilon_j$$

真实直方图  
第  $j$  区的探  
测效率



有时，没有真实事例发生，但观测到事例: 本底




$$v_i = \sum_{j=1}^M R_{ij} \mu_j + \beta_i$$

$\beta_i$  是在观测直方图上预期的本底数目。



# 各关键量汇总

真实直方图:  $\vec{\mu} = (\mu_1, \dots, \mu_M)$ ,  $\vec{\mu}_{\text{tot}} = \sum_{j=1}^M \mu_j$    $M$  个区间

概率:  $\vec{p} = (p_1, \dots, p_M) = \vec{\mu} / \mu_{\text{tot}}$    $N$  个区间

观测直方图的期待值:  $\vec{v} = (v_1, \dots, v_N)$

观测直方图:  $\vec{n} = (n_1, \dots, n_N)$

响应矩阵:  $R_{ij} = P(\text{观测值在第 } i \text{ 区} | \text{真实值在第 } j \text{ 区})$

效率:  $\varepsilon_j = \sum_{i=1}^N R_{ij}$  预期本底:  $\vec{\beta} = (\beta_1, \dots, \beta_N)$



$$E[\vec{n}] = \vec{v} = R\vec{\mu} + \vec{\beta}$$

一般通过构造  $\ln L$  或  $\chi^2$  寻求  $\vec{\mu}$  的估计量, 这需要相关的概率理论, 例如: 泊松分布(各区间独立)

$$P(n_i; v_i) = \frac{v_i^{n_i}}{n_i!} e^{-v_i}$$

或关联矩阵(各区间不独立)  $V_{ij} = \text{cov}[n_i, n_j]$

# 为什么要用解谱法？

---

一般而言，我们不需要解谱法，例如当比较现有理论的预期值时，最好是将探测器相应叠加到理论中去，即在预期值中包含探测器效应并与未修正的原始数据  $\vec{n}$  相比较。

但是，不将实验数据进行解谱处理，结果发表后，有关反应矩阵的知识将不再保留。而解谱后的分布可直接与各种理论的预言比较，也可与别的实验经过解谱以后的分布相比较。

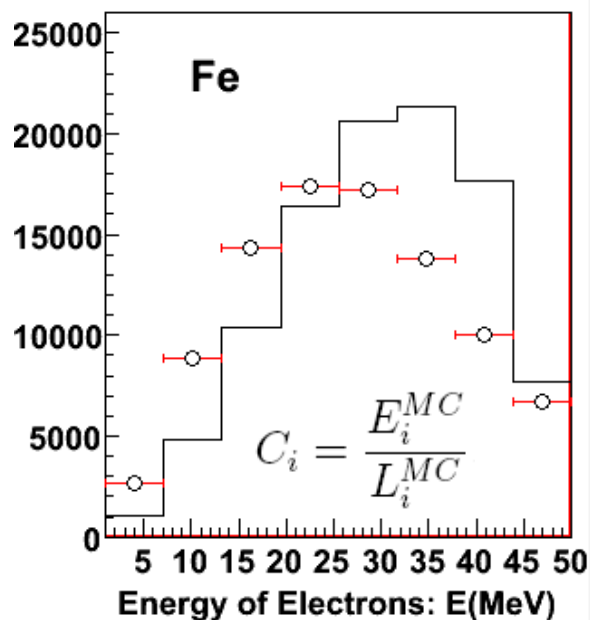
通常解谱后的结果更有用，否则当反应矩阵不可恢复时，即使对结果又有新的理论解释，也很难进行理论检验。

在粒子物理研究中，解谱法常用的领域为：

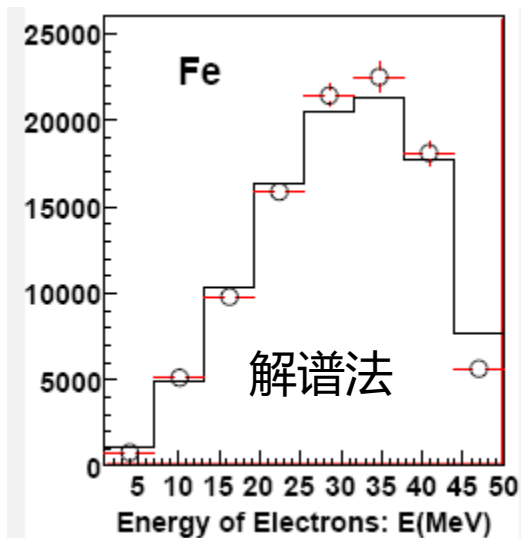
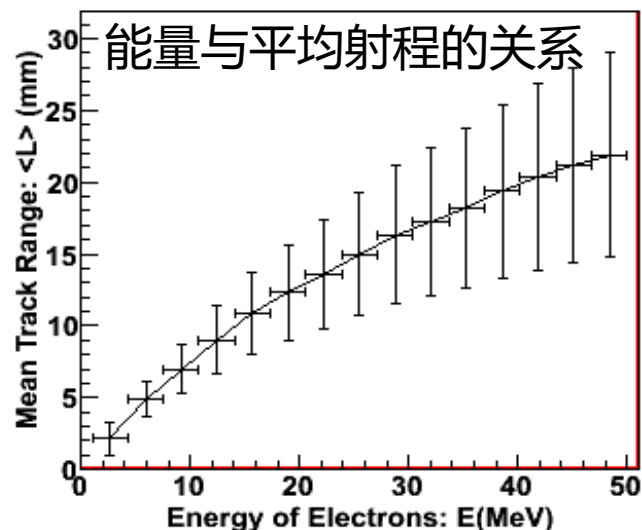
- 强子结构函数
- $\tau$  的谱函数(即强子质量谱)
- 强子事例形状分布
- 粒子多重数分布
- . . .

# 为什么用解谱法：举例

中微子与铁原子核相互作用，产生电子，测量电子的能谱。实验上只可能测量电子在铁中的射程。



因子修正法：能谱(圆圈)与真值(直方图)



# 响应矩阵的逆

假设  $\vec{v} = R\vec{\mu} + \vec{\beta}$  的逆存在:  $\vec{\mu} = R^{-1}(\vec{v} - \vec{\beta})$ 。

若数据服从泊松分布

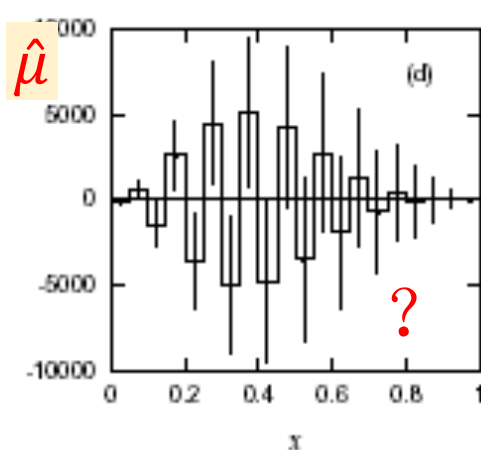
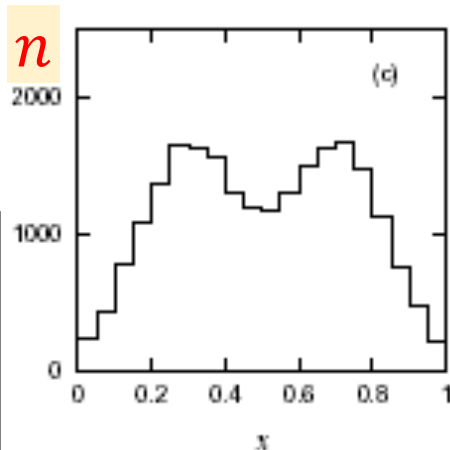
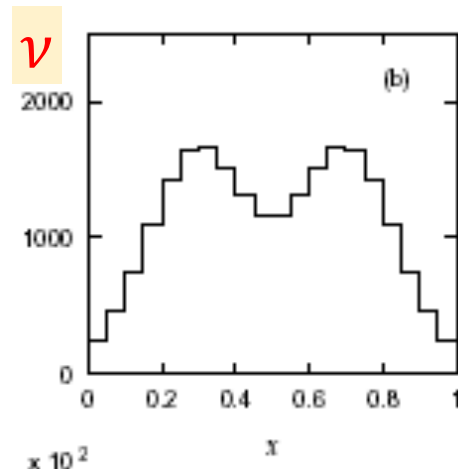
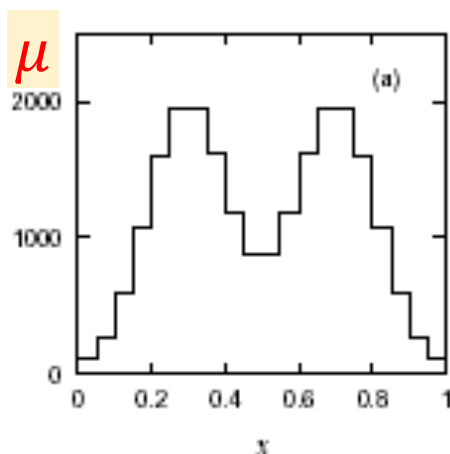
$$P(n_i; v_i) = \frac{v_i^{n_i}}{n_i!} e^{-v_i}$$

$$\rightarrow \ln L(\vec{\mu}) = \sum_{i=1}^N (n_i \ln v_i - v_i)$$

最大似然估计量为

$$\hat{\vec{v}} = \vec{n}, \quad \hat{\vec{\mu}} = R^{-1}(\vec{n} - \vec{\beta})$$

若  $R$  的非对角元太大, 即区间宽度比分辨率要小时, 会导致上式有很大的方差, 并且在相邻区间存在很强的负相关性。



# 估计量剧烈振荡的原因(I)

考虑一个简单的例子

$\varepsilon \rightarrow 1$ : 完美;  $\varepsilon \rightarrow 0$ : 分辨率很差

$$\hat{A}x = b, \text{ 其中 } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

$$\hat{A} = \frac{1}{2} \begin{pmatrix} 1 + \varepsilon & 1 - \varepsilon \\ 1 - \varepsilon & 1 + \varepsilon \end{pmatrix}, 0 < \varepsilon < 1$$

$$\begin{aligned} \hat{A}^{-1} &= \frac{1}{2\varepsilon} \begin{pmatrix} 1 + \varepsilon & -1 + \varepsilon \\ -1 + \varepsilon & 1 + \varepsilon \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1}{2\varepsilon} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{aligned}$$

$$x = \hat{A}^{-1}b = \frac{b_1 + b_2}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{b_1 - b_2}{2\varepsilon} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$\varepsilon \rightarrow 0$ 时, **第二项**起决定性作用。

当  $b_1 \simeq b_2$  时, 考虑到统计涨落,  $b_1 - b_2$  是中心值在零附近的随机数。这个随机数被  $1/\varepsilon$  放大, 结果中有用的信息完全被非物理的振荡湮没。

通常情况下, 直接求响应矩阵的逆得到真值  $x$  的方法, 尽管理论上是严格的, 且是有效估计量, 但结果没有物理意义。

解决办法是进行平滑处理, 消除无意义的统计涨落。  
但平滑会带来偏向性, 需要在涨落与偏向性之间找到平衡。

# 估计量剧烈振荡的原因(II)

考虑分布函数  $f(x) = \frac{a_0}{2} + \sum_{v=1}^{\infty} (a_v \cos vx + b_v \sin vx)$   
若真值 $x$ 的测量值按 $N(x, \sigma^2)$ 弥散, 则测量值 $y$ 的分布为

$$g(y) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y-x)^2}{2\sigma^2}\right] f(x) dx \quad (\text{卷积})$$

将 $g(y)$ 按  $\cos vy$  和  $\sin vy$  展开:

$$g(y) = \frac{\alpha_0}{2} + \sum_{v=1}^{\infty} (\alpha_v \cos vy + \beta_v \sin vy)$$

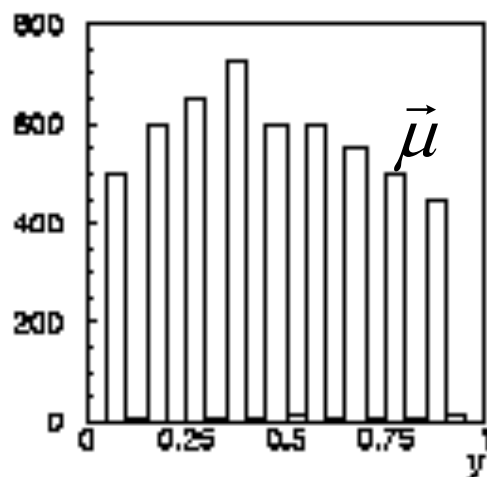
其中  $\alpha_v = \exp\left(\frac{v^2\sigma^2}{2}\right) \alpha_v$ ,  $b_v = \exp\left(\frac{v^2\sigma^2}{2}\right) \beta_v$

实际上, 非物理的剧烈涨落由高频部分主导。平滑处理主要是消除或压低高频部分无意义的涨落。

若可准确获得  $g(y)$  的系数  $\alpha_v$  和  $\beta_v$ , 则可严格求出真实分布  $f(x)$  的系数  $a_v$  和  $b_v$ 。但  $\alpha_v$  和  $\beta_v$  不可避免有统计涨落, 该涨落被放大  $\exp(v^2\sigma^2/2)$  倍。尤其对一般函数,  $v \rightarrow \infty$  时,  $a_v, b_v \rightarrow 0$ ,  $\alpha_v$  和  $\beta_v$  中的统计涨落被指数放大后将占据统治地位。

# 估计量剧烈振荡的原因(III)

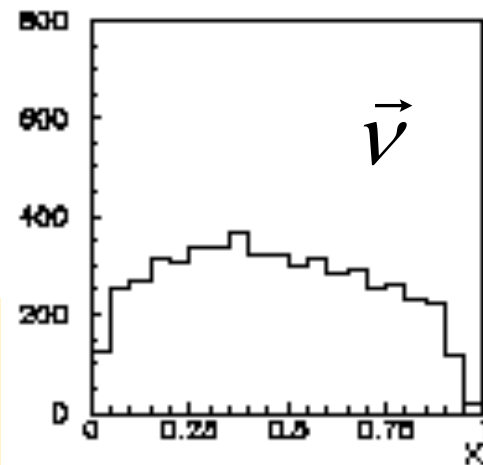
假设  $\vec{\mu}$  真存在精细结构



$R$ 作用后(理论上)  
得到  $\vec{v} = R\vec{\mu}$



大部分精细结构被抹平,  
但精细结构信息被细微地  
隐藏于 $\vec{v}$ 中。



$R^{-1}$  作用于  $\vec{v}$  应完全恢复精细结构:  $\vec{\mu} = R^{-1}\vec{v}$

但我们无法得到  $\vec{v}$ , 只有存在统计涨落的  $\vec{n}$ 。

$R^{-1}$ 无法区分 $\vec{n}$ 中的统计涨落和 $\vec{v}$ 中隐藏的精细结构, 认为统计涨落信息是 $\vec{\mu}$ 的精细结构造成, 从而恢复这种精细结构, 造成剧烈振荡效应。

# 重新理解最大似然解

---

估计量的均值

$$E[\hat{\vec{\mu}}] = R^{-1}(E[\vec{n}] - \vec{\beta}) = \vec{\mu} \quad \longrightarrow \quad \text{无偏!}$$

估计量的方差

$$\begin{aligned} U_{ij} = \text{cov}[\hat{\mu}_i, \hat{\mu}_j] &= \sum_{k,l=1}^N (R^{-1})_{ik} (R^{-1})_{jl} \text{cov}[n_k, n_l] \\ &= \sum_{k=1}^N (R^{-1})_{ik} (R^{-1})_{jk} v_k \end{aligned}$$

假设  $n_i$  是独立的泊松变量时,  $\text{cov}[n_k, n_l] = \delta_{kl} v_k$



# 重新理解最大似然解 (续)

利用 RCF 边界:

$$(U^{-1})_{kl} = -E \left[ \frac{\partial^2 \ln L}{\partial \mu_k \partial \mu_l} \right] = \sum_{i=1}^N \frac{R_{ik} R_{il}}{v_i}$$

$$\begin{aligned} \vec{\mu} &= R^{-1}(\vec{v} - \vec{\beta}) \\ \ln L(\vec{\mu}) &= \sum_{i=1}^N (n_i \ln v_i - v_i) \end{aligned}$$

求倒数得

$$U_{ij} = \sum_{k=1}^N (R^{-1})_{ik} (R^{-1})_{jk} v_k$$



与最大似然估计量的结果一致。

最大似然估计量在所有无偏估计中给出的方差最小。

但这个方差太大了!

为了减小方差, 必须引入一定的偏倚。

策略: 接受小的偏倚(系统不确定度)以换取大幅减小方差(统计不确定度)。

# 简单方法：修正因子法

---

对  $\vec{\mu}, \vec{v}$  做相同分区, 并取

$$\hat{\mu}_i = C_i(n_i - \beta_i)$$

$$C_i \equiv \frac{\mu_i^{\text{MC}}}{v_i^{\text{MC}}} \quad (\text{修正因子})$$

相当于  $R^{-1}$  取为对角矩阵:  
 $(R^{-1})_{ii} = C_i$

$v_i^{\text{MC}}$  与  $\mu_i^{\text{MC}}$  用蒙特卡罗模拟得到 (无本底)。

$$U_{ij} = \text{cov}[\hat{\mu}_i, \hat{\mu}_j] = C_i^2 \text{cov}[n_i, n_j]$$

通常  $C_i \approx O(1)$ , 因此方差不会被放大。

# 修正因子法中的偏倚

修正因子法的偏倚:  $b_i = E[\hat{\mu}_i] - \mu_i$

$$b_i = \left( \frac{\mu_i^{\text{MC}}}{v_i^{\text{MC}}} - \frac{\mu_i}{v_i^{\text{sig}}} \right) v_i^{\text{sig}}, \quad \text{其中 } v_i^{\text{sig}} = v_i - \beta_i = \varepsilon_i \mu_i$$

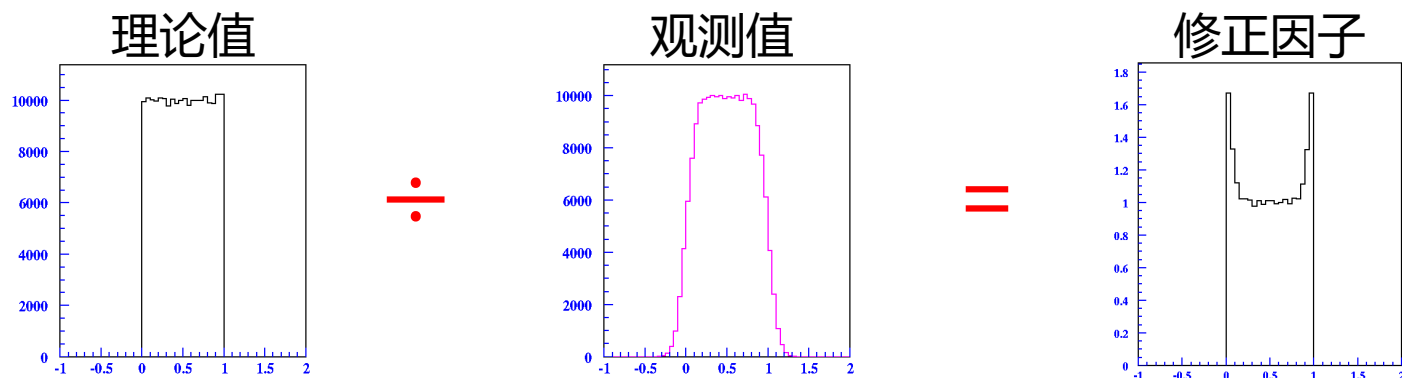
除非模拟采用的模型无误, 使得  $\mu_i^{\text{MC}} = \mu_i$ , 否则上式不为零, 需要考虑对应的系统不确定度。

注意: 该偏倚倾向于把  $\hat{\mu}$  拉向  $\vec{\mu}^{\text{MC}}$ , 造成模型检验的困难。

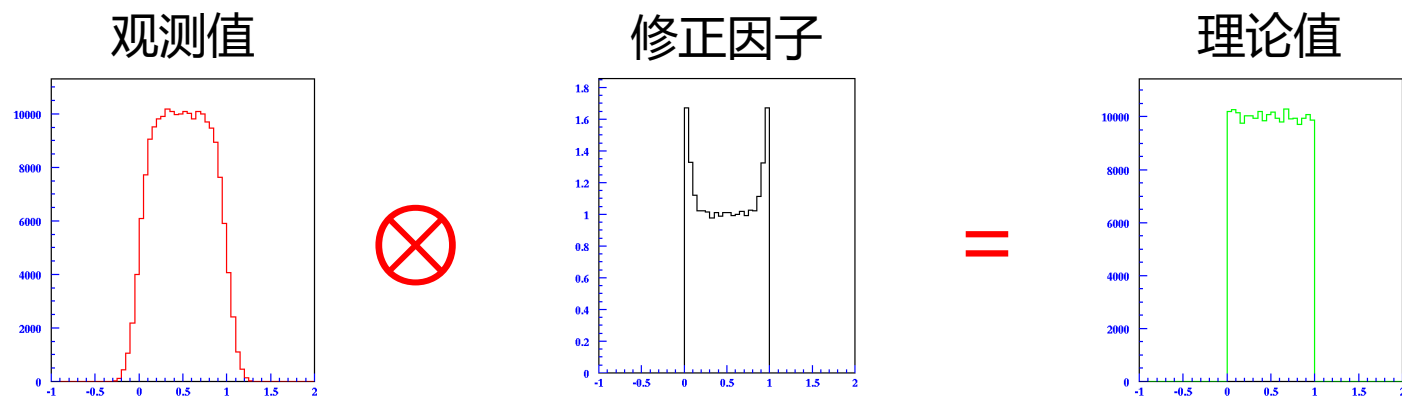
- 1) 如果分区宽度大于等于几倍分辨率, 结果不会太坏
- 2) 实际应用中, 该方法常用于事例形状变量的分布研究

# 例：脉冲形状的还原

理论(真实)直方图除以受实验仪器影响的直方图得到修正因子



观测直方图乘以修正因子直方图得到理论(真实)直方图



# 正规化的解谱法

---

考虑“合理的”估计量，对选定的  $\Delta \ln L$  满足

$$\ln L(\vec{\mu}) \geq \ln L_{\max} - \Delta \ln L$$

$\Delta \ln L$  描述了数据  $\vec{n}$  与期待值  $\vec{v}$  之间的“距离”。

$\vec{\mu}$  的估计量满足该不等式且最光滑，等价于求下式的最大值

$$\Phi(\vec{\mu}) = \alpha \ln L(\vec{\mu}) + S(\vec{\mu})$$

$S(\vec{\mu})$  = 正则化函数 (光滑性的量度)

$\alpha$  = 正则化参量 (其选择与给定的  $\Delta \ln L$  对应)

# 正规化的解谱法（续）

另外，要求解谱后对总事例数的估计为无偏的

$$\sum_{i=1}^N v_i = \sum_{i,j} R_{ij} \mu_j = n_{\text{tot}}$$

在约束情况下求下式最大值

$$\varphi(\vec{\mu}, \lambda) = \alpha \ln L(\vec{\mu}) + S(\vec{\mu}) + \lambda \left[ n_{\text{tot}} - \sum_{i=1}^N v_i \right]$$

因为  $\vec{v} = R\vec{\mu} + \vec{\beta}$ ,  
所以  $\vec{v}$  是  $\vec{\mu}$  的函数

$\lambda$ : 拉格朗日乘子

$$\frac{\partial \varphi}{\partial \lambda} = 0 \quad \Rightarrow \quad \sum_{i=1}^N v_i = n_{\text{tot}}$$

# 正规化的解谱法（续）

---

$$\varphi(\vec{\mu}, \lambda) = \alpha \ln L(\vec{\mu}) + S(\vec{\mu}) + \lambda \left[ n_{\text{tot}} - \sum_{i=1}^N v_i \right]$$

$\left\{ \begin{array}{l} \alpha = 0 : \text{给出最光滑的解(与数据无关)} \\ \alpha \rightarrow \infty : \text{给出最大似然解(方差太大)} \end{array} \right.$

需要正规化函数  $S(\vec{\mu})$  以及选取  $\alpha$  值的方案。

- a) Tikhonov 规则
- b) 最大熵(MaxEnt) 规则

不同方案得到的估计量的好坏由它们的偏倚和方差来判断。

# Tikhonov 规则

---

取光滑度等于第  $k$  阶导数平方的均值，有

$$S[f_{\text{true}}(y)] = - \int \left( \frac{d^k f_{\text{true}}(y)}{dy^k} \right)^2 dy, \quad \text{其中 } k = 1, 2, \dots$$

通常取  $k = 2$ ，使得  $S$  约等于曲率平方的平均值。  
对直方图而言，也就是

$$S[\vec{\mu}] = - \sum_{i=1}^{M-2} (-\mu_i + 2\mu_{i+1} - \mu_{i+2})^2$$

Sov. Math.5(1963)1035

**注意：**二阶导数对直方图的首尾两个区间没有好的定义。



# Tikhonov 规则 (续)

---

如果在  $\ln L = -\frac{1}{2}\chi^2$  条件下, 采用Tikhonov( $k = 2$ )规则

$$\varphi(\vec{\mu}, \lambda) = \alpha \ln L(\vec{\mu}) + S(\vec{\mu}) = -\frac{\alpha}{2}\chi^2(\vec{\mu}) + S(\vec{\mu})$$

这是  $\mu_i$  的二次项。对  $\varphi$  求偏微分, 给出线性方程, 得到  $\mu_i$  的估计值与方差。

在高能物理界现有好几个现成的程序:  
RUN, Blobel, SVD, Höcker, ...

# 最大熵(MaxEnt)规则

---

另一种表征光滑度的方法基于熵。

对于一组概率而言，熵定义为

$$H = - \sum_{i=1}^M p_i \ln p_i$$

$$p_i \equiv \frac{\mu_i}{\mu_{\text{tot}}}$$

Ann. Rev. Astron. Astrophys.24 (1986)127

所有  $p_i$  相等意味着熵最大(最光滑)。

某个  $p_i = 1$ , 其它都为零, 则意味着熵最小。

# 最大熵(MaxEnt)规则 (续)

用熵作为正规化函数:

$$S(\vec{\mu}) = H(\vec{\mu}) = - \sum_{i=1}^M \frac{\mu_i}{\mu_{\text{tot}}} \ln \frac{\mu_i}{\mu_{\text{tot}}} \\ \propto \ln(\mu_{\text{tot}} \text{填入} M \text{个区间的各种可能方式的总数})$$

有时, 根据贝叶斯统计

$$S(\vec{\mu}) \rightarrow \vec{\mu} \text{的先验概率密度函数(?)}$$

这里, 我们仍采用经典近似: 估计量的好坏由偏倚和方差来判断。

注意: 熵与区间的顺序无关。

# $\hat{\vec{\mu}}$ 的方差与偏倚

一般来说, 决定  $\hat{\vec{\mu}}(\vec{n})$  的方程是非线性的。在得到正规函数后, 将  $\hat{\vec{\mu}}(\vec{n})$  在  $\vec{n}_{\text{obs}}$  附近展开:

$$\hat{\vec{\mu}}(\vec{n}) \approx \hat{\vec{\mu}}_{\text{obs}} - A^{-1}B(\vec{n} - \vec{n}_{\text{obs}})$$

$$A_{ij} = \begin{cases} \frac{\partial^2 \varphi}{\partial \mu_i \partial \mu_j}, & i, j = 1, \dots, M, \\ \frac{\partial^2 \varphi}{\partial \mu_i \partial \lambda} = -1, & i = 1, \dots, M; j = M + 1, \\ \frac{\partial^2 \varphi}{\partial \lambda^2} = 0, & i = M + 1; j = M + 1, \end{cases}$$

$\varphi$  是非正规的似然函数

$$B_{ij} = \begin{cases} \frac{\partial^2 \varphi}{\partial \mu_i \partial n_j}, & i = 1, \dots, M; j = 1, \dots, N, \\ \frac{\partial^2 \varphi}{\partial \lambda \partial n_j} = 1, & i = M + 1; j = 1, \dots, N. \end{cases}$$

G. Cowan, *Statistical Data Analysis*,  
Oxford University Press (1998).

# $\hat{\vec{\mu}}$ 的方差与偏倚 (续)

---

利用误差传递公式得到协方差:  $U_{ij} = \text{cov}[\hat{\mu}_i, \hat{\mu}_j]$

$$U = CV C^T$$

其中:  $C = A^{-1}B$ ,  $V_{ij} = \text{cov}[n_i, n_j]$

偏倚  $b_i = E[\hat{\mu}_i] - \mu_i$  的估计量:

$$\hat{b}_i = \sum_{j=1}^N C_{ij}(\hat{v}_j - n_j) = \sum_{j=1}^N \frac{\partial \hat{\mu}_i}{\partial n_j}(\hat{v}_j - n_j)$$

此处  $\hat{\vec{v}} = R\hat{\vec{\mu}} + \vec{\beta}$ , 而且通常情况下  $\hat{\vec{v}} \neq \vec{n}$ 。

# 正规化参数 $\alpha$ 的选取

$\alpha$  决定了赋予数据的权重以便能与光滑度相比较。

$\alpha \rightarrow 0$  : 给出光滑度最大的估计值, 并与数据无关;  
虽然方差为零, 但偏倚明显。

$\alpha \rightarrow \infty$ : 回归无偏但高度振荡的最大似然解。

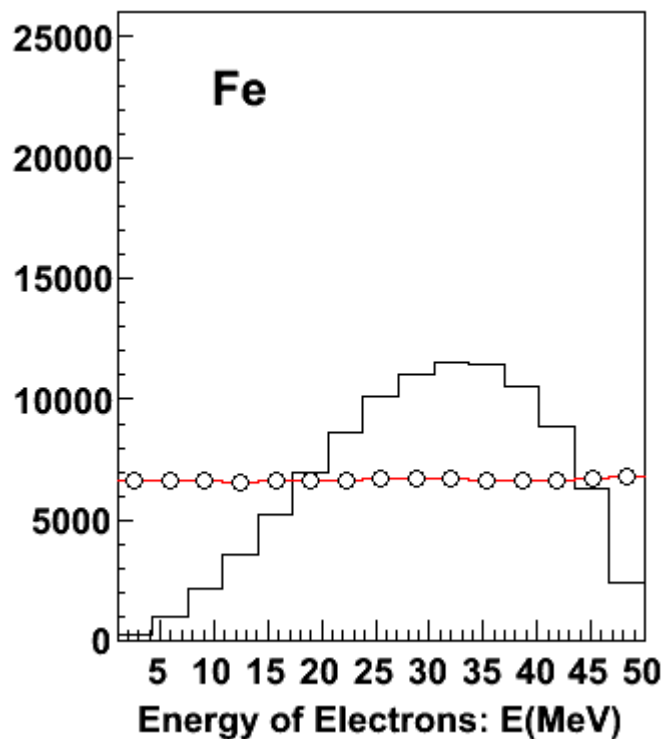
偏倚与方差之间的平衡: 选择  $\alpha$  使均方误差 (MSE) 最小

$$\text{MSE} = \frac{1}{M} \sum_{i=1}^M (U_{ii} + \hat{b}_i^2), \quad \text{或加权的} \quad \text{MSE} = \frac{1}{M} \sum_{i=1}^M \frac{U_{ii} + \hat{b}_i^2}{\hat{\mu}_i}$$

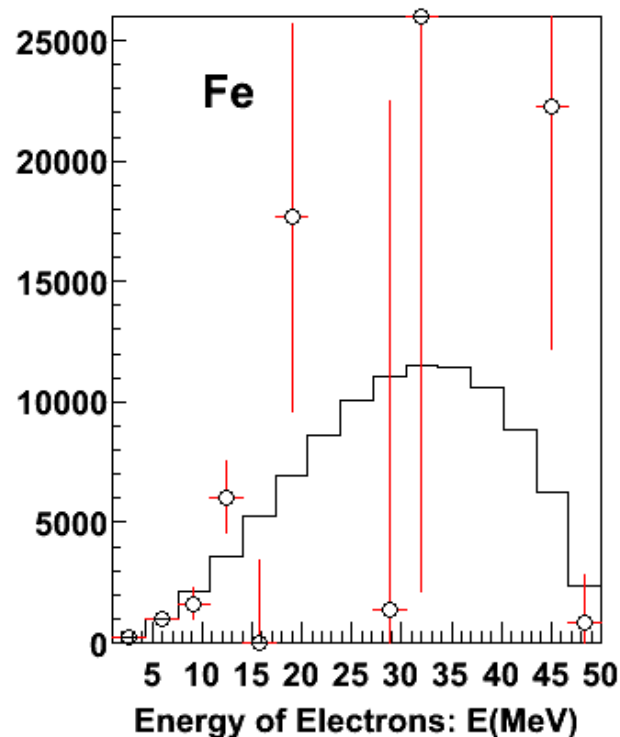
或要求偏倚不大于它自身的估计方差  $\hat{W}_{ii}$ 。找到  $\alpha$  的值使得

$$\chi_b^2 = \sum_{i=1}^M \frac{\hat{b}_i^2}{\hat{W}_{ii}} = M \quad \text{其中} \quad \hat{W}_{ij} = \text{cov}[\hat{b}_i, \hat{b}_j]$$

# 正规化参数选择的重要性



$\alpha \rightarrow 0$  : 光滑,  
完全由MC决定,  
与数据无关



$\alpha \rightarrow \infty$  : 完全由数据决定,  
剧烈振荡,  
无物理意义


# 贝叶斯迭代解谱法

$R_{ij} = P(\text{测量值在区间}i|\text{真值在区间}j)$

利用贝叶斯定理，可以得到估计量：

$$\begin{aligned}\hat{\mu}_i &= \frac{1}{\varepsilon_i} \sum_{j=1}^N n_j P(\text{真值在区间}i|\text{测量值在区间}j) \\ &= \frac{1}{\varepsilon_i} \sum_{j=1}^N n_j \frac{P(\text{测量值在区间}j|\text{真值在区间}i) p_i}{\sum_k P(\text{测量值在区间}j|\text{真值在区间}k) p_k} \\ &= \frac{1}{\varepsilon_i} \sum_{j=1}^N \frac{R_{ji} p_i}{\sum_k R_{jk} p_k} n_j\end{aligned}$$

应用贝叶斯定理



选取初始  $p_i (i = 1, \dots, N)$ ，之后用  $p_i = \hat{\mu}_i / \mu_{\text{tot}}$  迭代。  
用测试数据确定迭代次数，之后用于实验数据。  
当迭代次数很大时，结果区域剧烈振荡。



# SVD方法解谱

Nucl. Instrum. Meth. A372, 469 (1996)

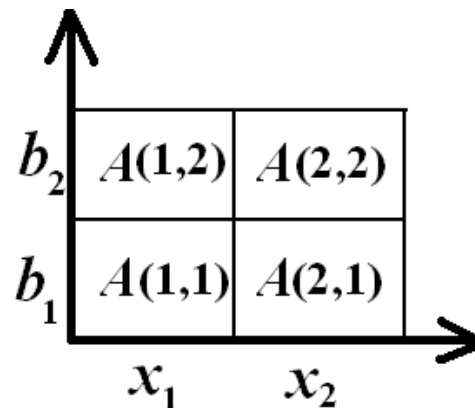
解谱需要先得到响应矩阵  $R$  并求逆。  
若  $R$  矩阵性质不太好，SVD方法更可靠。

用MC得到  $R$ ：

$x$ ：MC真值       $b$ ：模拟测量值

作 $x$ : $y$ 二维直方图（以 $x$ 和 $b$ 都分2个区间为例）

实际上得到一个  $2 \times 2$  的矩阵（或数组） $A$



$A(1,1)$ ：真值为  $x_1$ ，测量值为  $b_1$  的事例数  $\propto R_{11}$

$A(1,2)$ ：真值为  $x_1$ ，测量值为  $b_2$  的事例数  $\propto R_{21}$

$A(2,1)$ ：真值为  $x_2$ ，测量值为  $b_1$  的事例数  $\propto R_{12}$

$A(2,2)$ ：真值为  $x_2$ ，测量值为  $b_2$  的事例数  $\propto R_{22}$

$$R_{ij} = \frac{A(j, i)}{\sum_i A(j, i)}$$



$$R \propto A^T$$

# Singular Value Decomposition(SVD)

任意  $n \times m$  矩阵  $A$  可以分解为  $A = USV^T$   
其中  $U = AA^T$  的本征矢构成一个  $n \times n$  矩阵  
 $V = A^T A$  的本征矢构成一个  $m \times m$  矩阵

$S = \sqrt{\text{diag}(\text{eig}(AA^T))}$ :  $n \times m$  对角矩阵

$$Ax = b \Rightarrow USV^T x = b \Rightarrow SV^T x = U^T b$$

定义  $z = V^T x$ ,  $d = U^T b$

$$\Rightarrow Sz = d \Rightarrow z = S^{-1}d \Rightarrow x = Vz = VS^{-1}d = VS^{-1}U^T b = A^{-1}b$$

SVD把  $b$ 和 $x$  都分解成正交归一向量的线性组合, 其系数分别构成向量  $d$ 和 $z$ , 基分别是矩阵  $U$ 和 $V$ 的列。

SVD后, 已知量  $b$  变为  $d$ , 未知量  $x$  变为  $z$ , 它们通过对角矩阵  $S$  联系起来, 对角元素为矩阵  $A$  的奇异值。

奇异值很小的部分是结果剧烈振荡的根源 (参见2维例子中  $\varepsilon \rightarrow 0$  的情况), 对应于傅立叶分解中高频部分, 相应地, 向量  $d$  中相应的元素服从标准正态分布。等价地说, 低频部分是物理的, 高频部分需要截断, 从  $d$  的元素分布可以找到截断值。

$U, V$  满足  $UU^T = I$ ,  $VV^T = I$   
 $S_{ii} \geq S_{jj}$ , 如果  $i < j$ 。  
显然  $A^{-1} = VS^{-1}U^T$

# TSVDUnfold

---

ROOT提供了SVD解谱法:

**TSVDUnfold (const TH1D \*bdat, const TH1D \*bini, const TH1D \*xini, const TH2D \*Adet)**

TSVDUnfold \*tsvdunf = new TSVDUnfold( bdat, Bcov, bini, xini, Adet );

TH1D\* unfresult = tsvdunf->Unfold( kreg );

**Measured spectrum (to be unfolded):**

bdat with covariance matrix Bcov

**Monte Carlo inputs:**

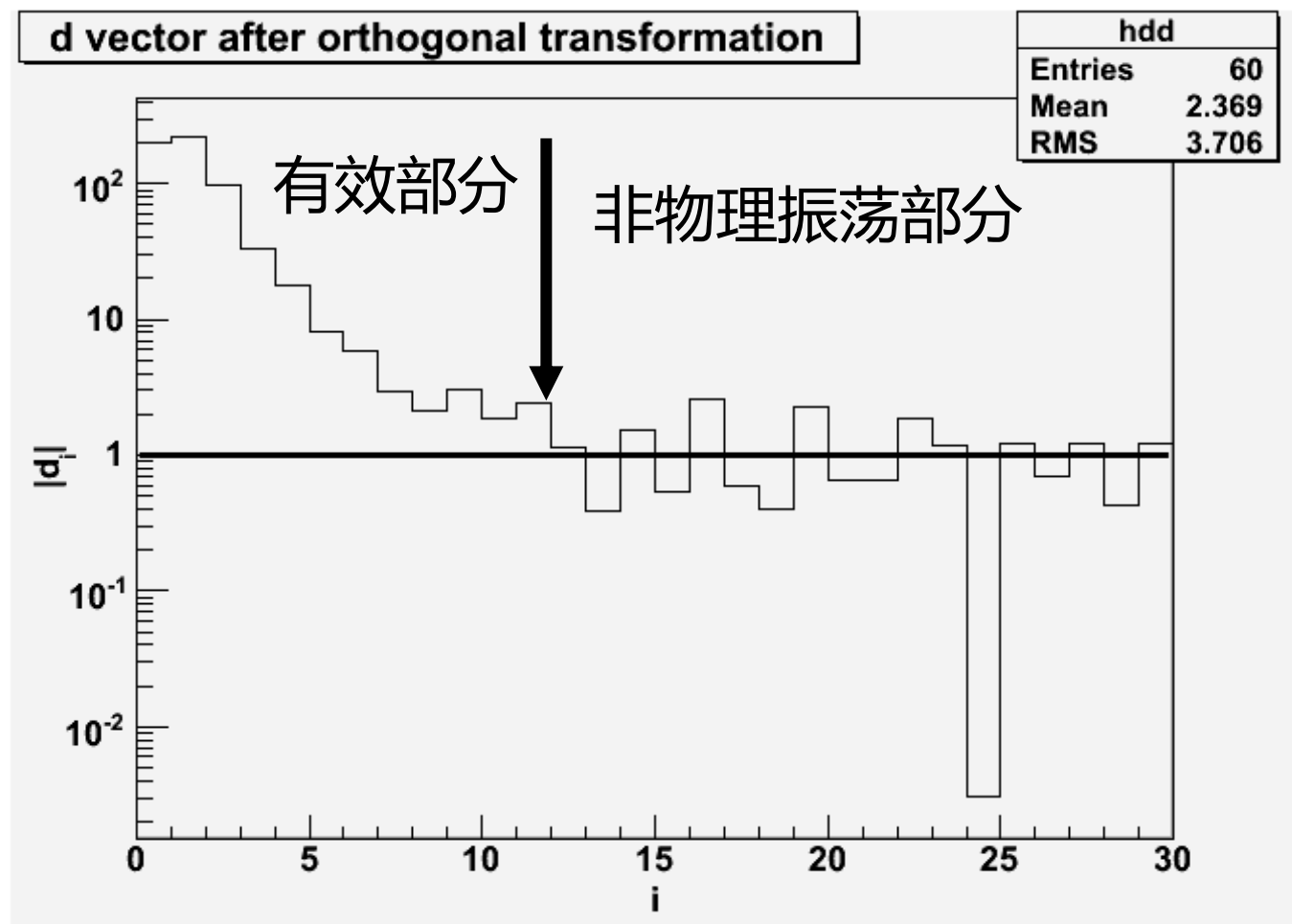
xini: true underlying spectrum (TH1D, n bins)

bini: reconstructed spectrum (TH1D, n bins)

Adet: response matrix (TH2D, nxn bins)

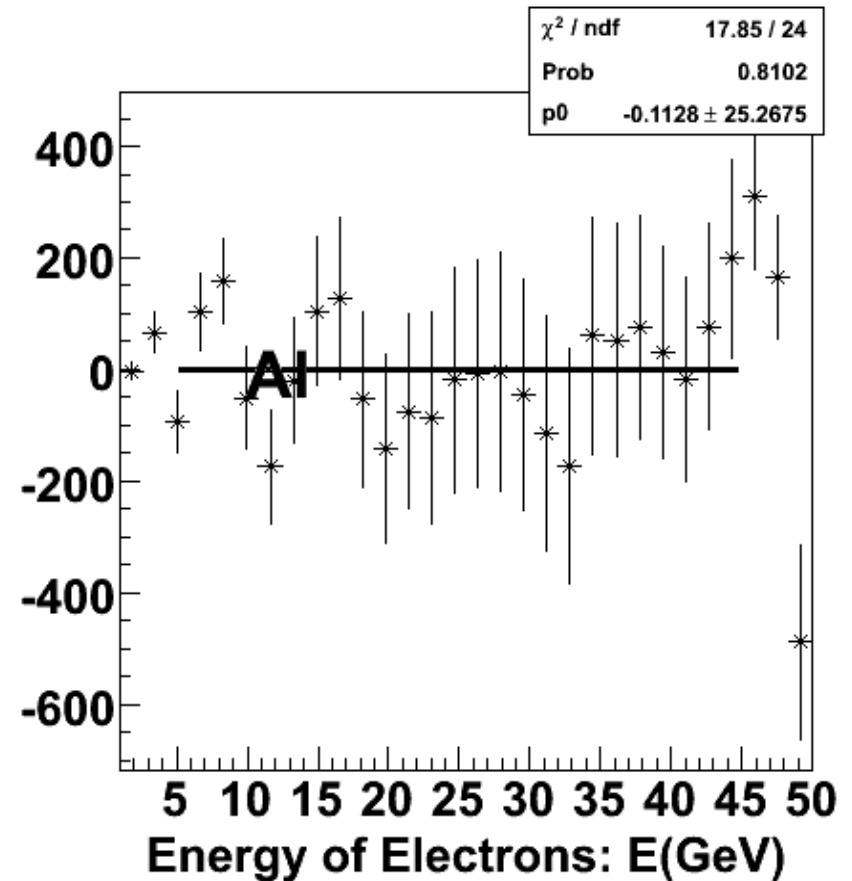
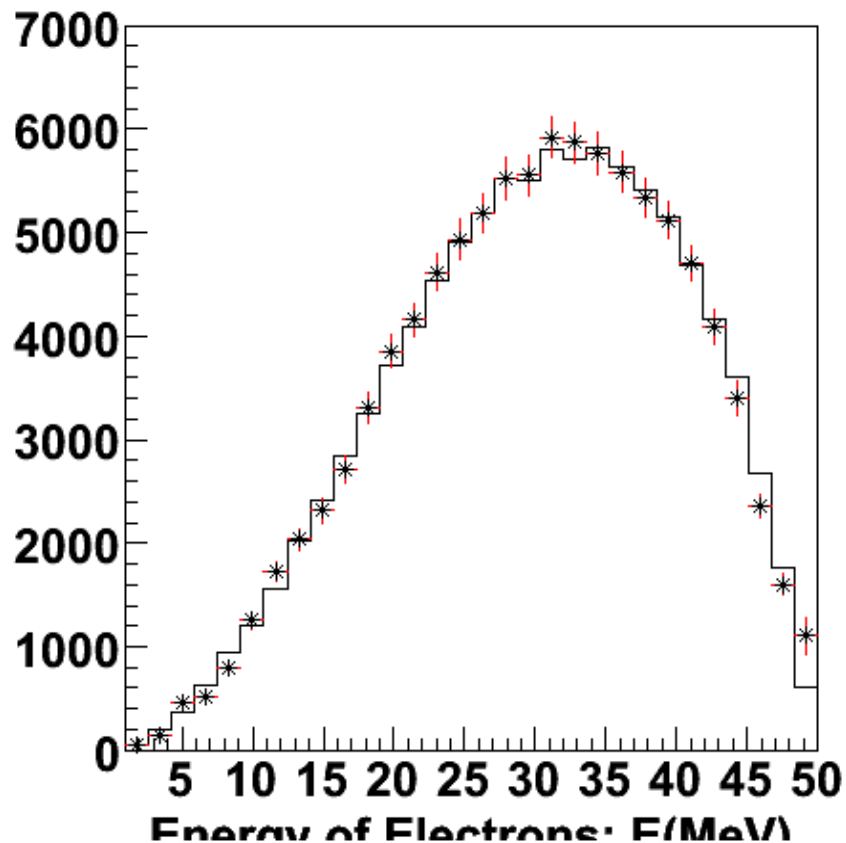
**kreg determines the regularisation of the unfolding.**

# 正规化参数选择的例子

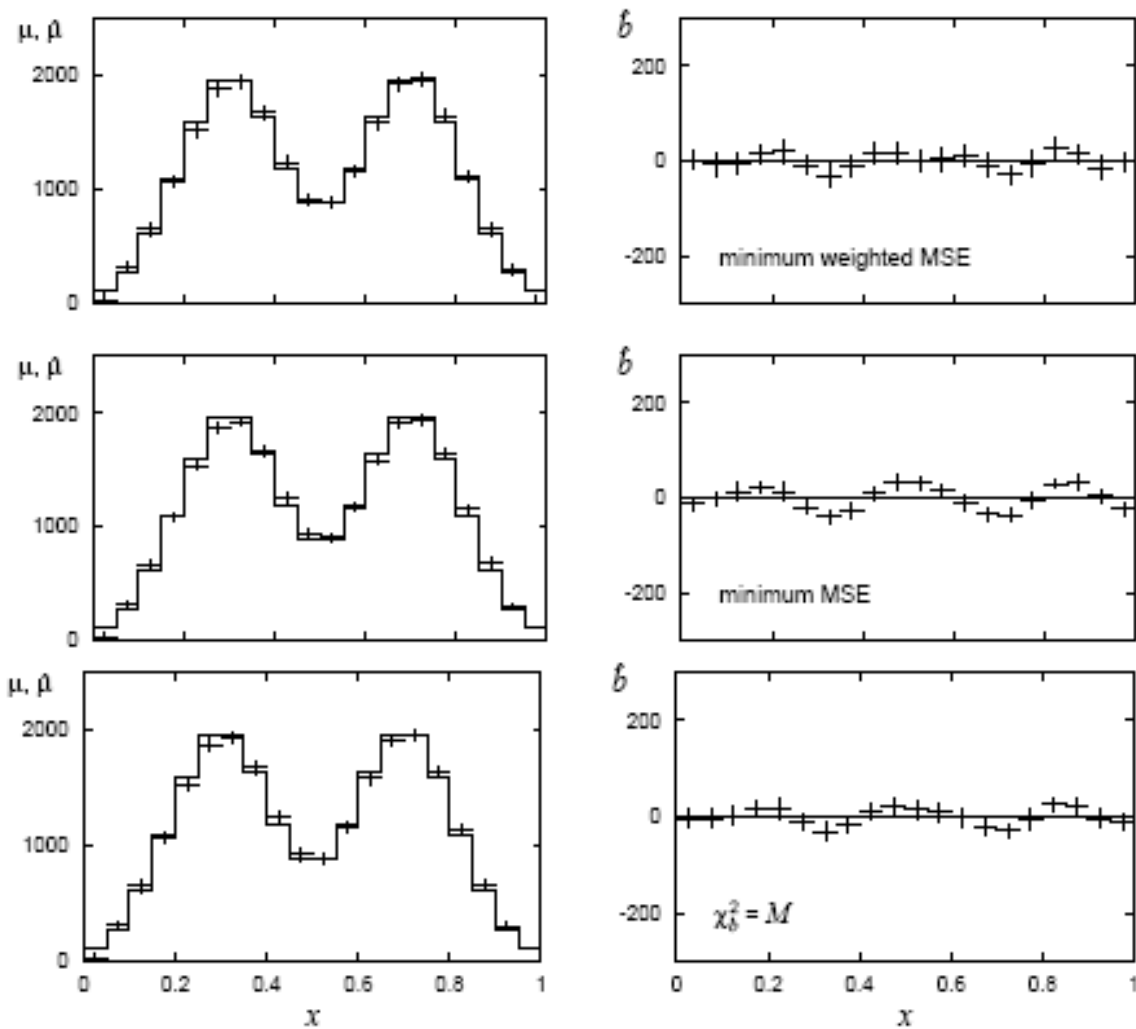


在该例子中，kterm选择为12左右。

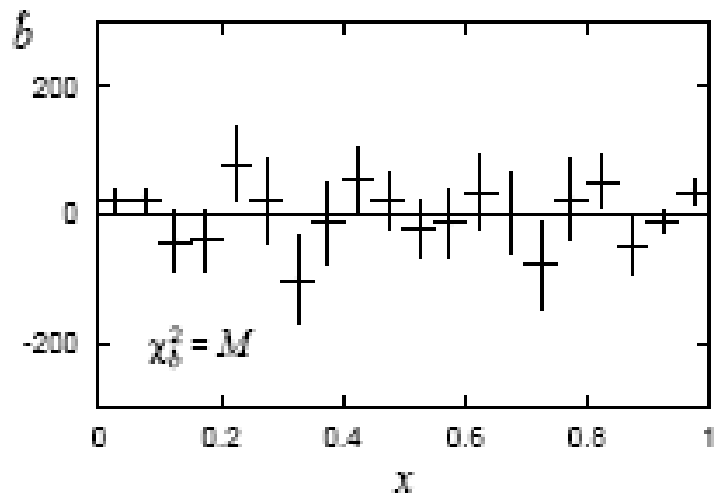
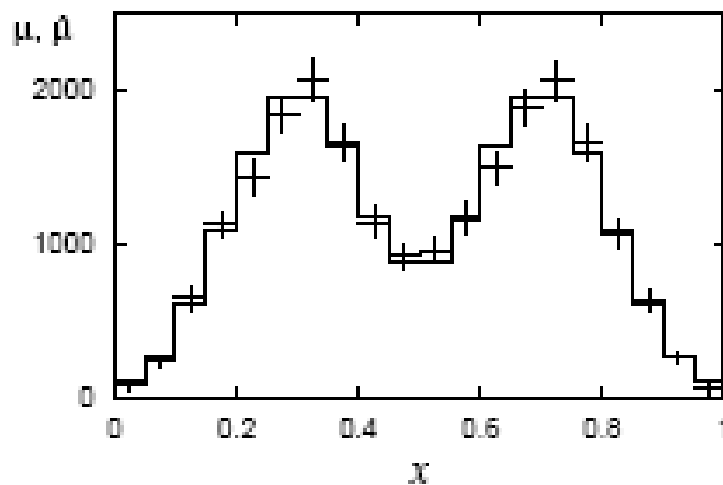
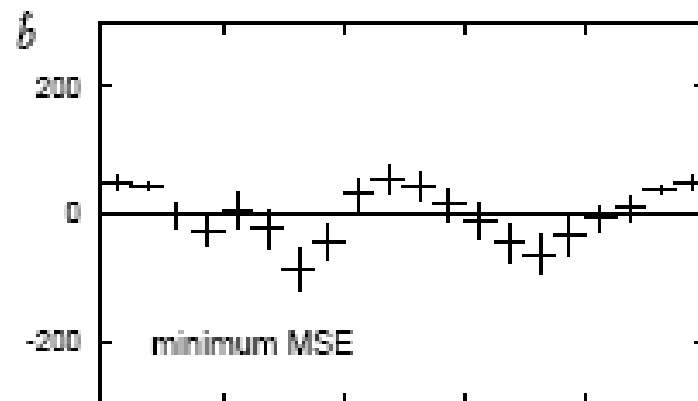
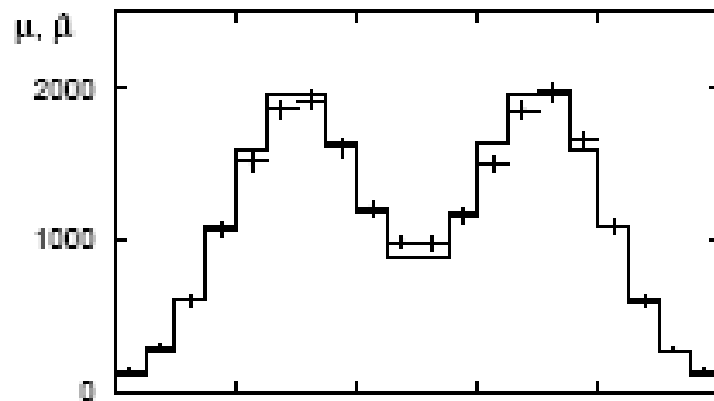
# 例: TSVDUnfold



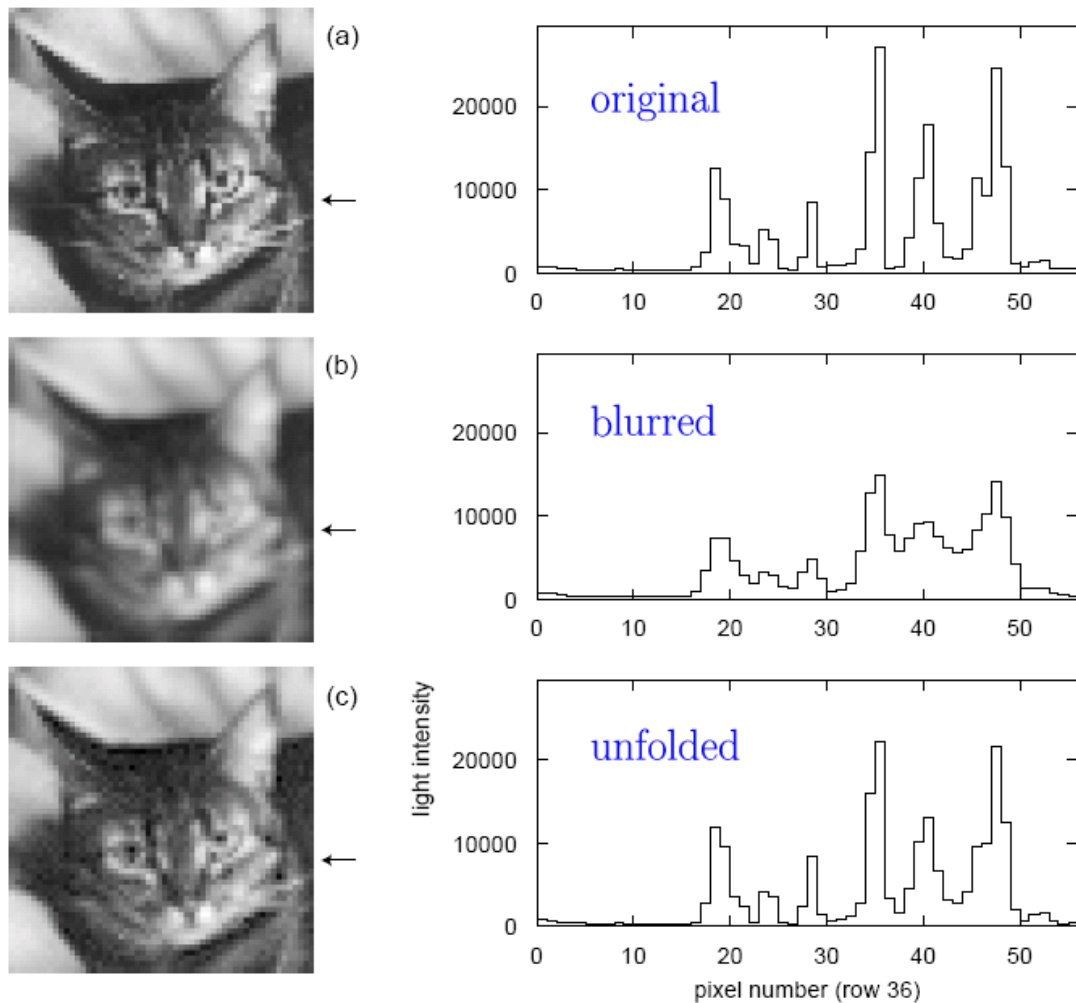
# 例: Tikhonov规则( $k=2$ )



# 例: 最大熵(MaxEnt)规则



# 例：图像处理中的最大熵规则



最大熵方法常用于天文观测图像的重建，与点源的偏倚较小，易于推广到二维以上的情况。



# 小结

---

## 1. 数学描述

真实直方图  $\vec{\mu}$ , 数据  $\vec{n}$  及其期待值  $\vec{v}$ , 满足  $\vec{\mu} = R\vec{v} + \vec{\beta}$ , 目标是构造  $\vec{\mu}$  的估计量。

## 2. 求反应矩阵的逆

有很大的振荡行为 (方差很大), 但却有效估计量。

## 3. 修正因子

$C_i = \mu_i^{\text{MC}} / v_i^{\text{MC}}$ , 快且简单

## 4. 正规化解谱

Tikhonov: 从第  $k$  阶导数的均方差中进行光滑处理  
MaxEnt : 从  $H = -\sum p_i \ln p_i$  熵中进行光滑处理

## 5. 估计量的方差与偏倚

求解过程中采用了线性近似, 因而不是无偏的

## 6. 正规化参数的选择

没有最佳方案, 可以采用  $\chi^2 = M$  (区间总数)

## 7. TSVDUnfold及解谱法例子

只要探测器的响应矩阵可知, 就一定可以得到真实分布