International Atomic Energy Agency

Evaluation: Least Squares and Weighted Mean

Naohiko Otsuka

IAEA Nuclear Data Section

Various Evaluation Methodology

Evaluation methodology depends on availability of experimental data and model:

- Eye-guide: Historical approach, *no longer acceptable*….
- Least-squares fitting to experimental data
  (e.g., fission cross sections for standards)
- Model calculation
- Mixture of experimental data and model calculation data
  (e.g., Unified Monte Carlo approach UMC–Capote+Smith).

Model and Model Parameter

Adjustment of model parameters **p** to reproduce experimental data **y** is often an essential part of data evaluation, for example

- Optical potential parameters to reproduce experimental (total) reaction cross sections.
- Fission barrier parameters to reproduce experimental fission cross sections.
- Fission neutron multiplicities to reproduce experimental k-eff.
- Evaluated cross section to reproduce experimental cross section.

Model parameters **p** and experimental data **y** are related by a model **f**: $\mathbf{y}=\mathbf{f}(\mathbf{p})+\mathbf{\Delta}$ (or $\mathbf{y}\sim\mathbf{f}(\mathbf{p})$).

The model **f** does not always have analytical expression. (e.g., input and output of reaction model codes).

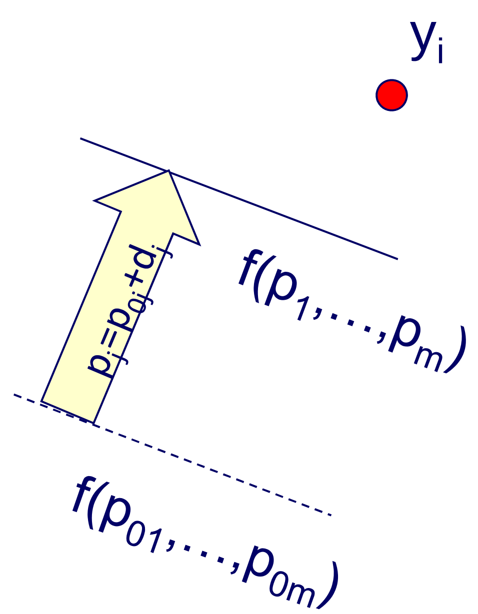
Conventional Least Squares

Problem:

1.There is an old model parameter set p_{0j} (j=1,m) which is related with experimental observable by the fitting function f: $f_i(p_{01},\ldots,p_{0m})$.

2.<u>New experimental data points</u> y_i (i=1,n) are available.

3.We want to have <u>an updated new model parameter set</u> p_j (j=1,m) by using new experimental data points y_i (i=1,n).



Minimize Chi-Square

Difference between in new and old model parameters:

$p_j=p_{0j}+d_j$ (j=1,m).

This makes the following change in $f_i(p_1,\ldots,p_m)$:

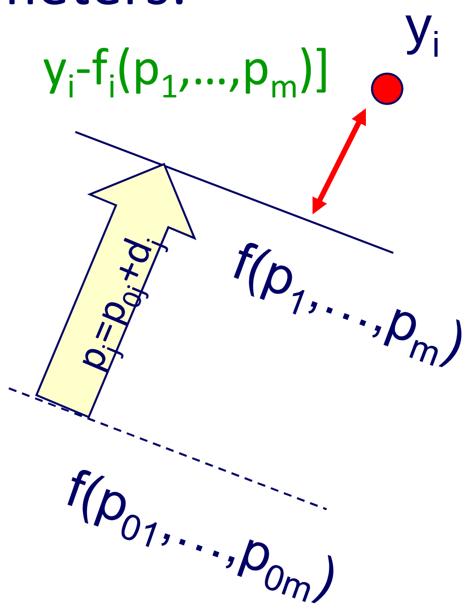
$f_i(p_1,\ldots p_m)\sim f_i(p_{01},\ldots,p_{0m})+\Sigma_{j=1,m}\, a_{ij}\cdot d_j$

with $a_{ij}=\partial f_i/\,\partial p_j\,|_{p=p0}$.

d_j must be chosen so that the sum of deviation between y_i and $f_i(p_1,\ldots,p_m)$ is minimum:

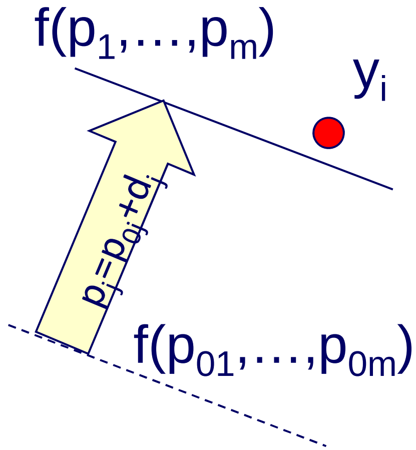
$\chi^2=Q(d_1,\ldots,d_m)=\Sigma_{i=1,n}[y_i-f_i(p_1,\ldots,p_m)]^2=\text{Min.}$

if we trust all experimental points by the same weight.



old mod. param. p_{0j} (j=1,m), updated mod. param. p_j (j=1,m)
new exp. points y_i (i=1,n), model f: y~f(p)

Minimum Condition Weighted by Variance



The minimum condition

$Q(d_1,\ldots,d_m)=\Sigma_{i=1,n}[y_i-f_i(p_1,\ldots,p_m)]^2=\text{Min.}$

is revised to

$Q(d_1,\ldots,d_m)=\Sigma_{i=1,n}[y_i-f_i(p_1,\ldots,p_m)]^2/(\Delta y_i)^2=\text{Min.}$

if each experimental point y_i has its standard deviation $\Delta y_i=(V_{ii})^{1/2}$.

This minimum condition gives m equations:

$\partial Q(d_1,\ldots,d_m)/\partial d_i=0\ (i=1,m)$.

old mod. param. p_{0j} (j=1,m), updated mod. param. p_j (j=1,m)
new exp. points y_i (i=1,n), model f: y~f(p)

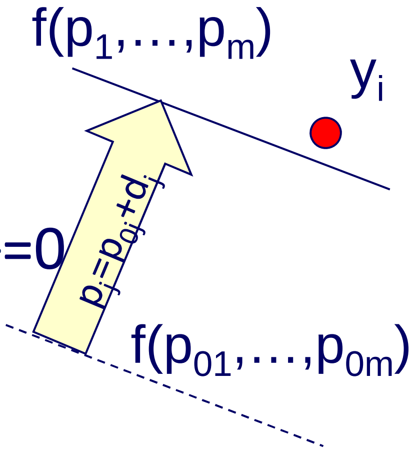
Equations for Parameter Shift $p - p_0$

$Q(d_1,\ldots,d_m)=\Sigma_{i=1,n}[y_i-f_i(p_1,\ldots,p_m)]^2/V_{ii}=\text{Min.}$

$\rightarrow \partial Q(d_1,\ldots,d_m)/\partial d_j=0 \quad (j=1,m).$

$\rightarrow \partial Q(d_1,\ldots,d_m)/\partial d_j=\{\Sigma_{i=1,n}2[y_i-f_i(p_1,\ldots,p_m)]/V_{ii}](\partial f_i/\partial d_j)\}=0$

$\{\Sigma_{i=1,n}[y_i-f_i(p_1,\ldots,p_m)]^2/V_{ii}]\cdot \underline{(\partial f_i/\partial d_j)}\}=0 \quad (j=1,m)$ –(*)



Recalling the linear expansion,

$f_i(p_1,\ldots p_m)\sim f_i(p_{01},\ldots,p_{0m})+\Sigma_{j=1,m}\, a_{ij}\cdot d_j$ with $a_{ij}=\partial f_i/\partial p_j|_{p=p0}$. $(j=1,m)$,

we obtain $(\partial f_i/\partial d_j)=a_{ij}$.

By putting this relation to (*),

$\Sigma_{i=1,n}[y_i- f_i(p_{01},\ldots,p_{0m})-\Sigma_{k=1,m}\, a_{ik}\cdot d_k\,]\, a_{ij}/V_{ii}=0 \ (j=1,m)$. (normal eq.)

old mod. param. p_{0j} $(j=1,m)$, updated mod. param. p_j $(j=1,m)$
new exp. points y_i $(i=1,n)$, model f: $y\sim f(p)$

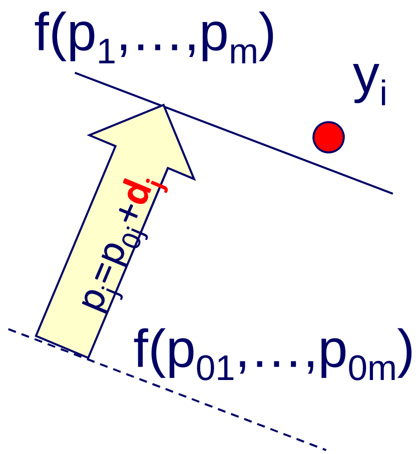
Normal Equation

$Q(d_1,\ldots,d_m)=\Sigma_{i=1,n}[y_i-f_i(p_1,\ldots,p_m)]^2/V_{ii}=\min.$

$\rightarrow$

$\Sigma_{i=1,n}[y_i- f_i(p_{01},\ldots,p_{0m})-\Sigma_{k=1,m}\, a_{ik}\cdot d_k\,]\, a_{ij}/V_{ii}=0 \quad (j=1,m).$

with $a_{ij}=\partial f_i/\,\partial p_j\,|_{p=p0}$.



We obtain m equations for d_j (j=1,m) – "Normal equation".

All y_i, V_{ii}, a_{ij}, p_{0j} are known with the analytical form of f_i (i=1,n; j=1,m), and we can solve these equations for d_j (j=1,m) which minimize $Q(d_1,\ldots,d_m)$.

old mod. param. p_{0j} (j=1,m), updated mod. param. p_j (j=1,m)
new exp. points y_i (i=1,n), model f: y~f(p)

Conventional Least Squares

$\Sigma_{i=1,n}(a_{ij}/V_{ii})[y_i - f_i(p_{01},\ldots,p_{0m}) - \Sigma_{k=1,m}\, a_{ik}\cdot d_k\,] = 0 \quad (j=1,m).$

with $a_{ij}=\partial f_i/\,\partial p_j\,|_{p=p0}$, $d_k=p_k-p_{0k}$

Let us introduce the following vector/matrix notation:

$\mathbf{y}=\{y_i\}$, $\mathbf{y_0}=\{f_i(p_{01},\ldots,p_{0m})\}$ (n dimensional vectors)

$\mathbf{p}=\{p_j\}$, $\mathbf{p_0}=\{p_{0j}\}$ (m dimensional vectors)

$\mathbf{G}=\{a_{ij}\}$ (n × m matrix), $\mathbf{V}^{-1}=\{1/V_{ii}\}$ (n × n diagonal matrix)

, then $\mathbf{G}^t\mathbf{V}^{-1}[\mathbf{y}-\mathbf{y_0}-\mathbf{G}(\mathbf{p}-\mathbf{p_0})]=0.$

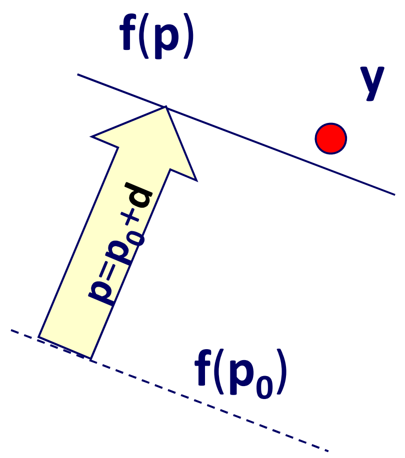
This gives the updated fitting parameters $\mathbf{p}=\{p_j\}$:

$$\mathbf{p} = \mathbf{p_0} + (\mathbf{G}^t\mathbf{V}^{-1}\mathbf{G})^{-1}\mathbf{G}^t\mathbf{V}^{-1}(\mathbf{y}-\mathbf{y_0})$$

old mod. param. p_{0j} (j=1,m), updated mod. param. p_j (j=1,m)
new exp. points y_i (i=1,n), model f: y~f(p)

What We Solved??

1. Default parameter $\mathbf{p_0}$ for model **f**
2. New experimental data **y**
3. The parameter $\mathbf{p_0}$ describe **y** by $\mathbf{f(p_0)}$
4. Find **p** which minimize "distance" between **y** and **f(p)**.



$$\rightarrow p = p_0 + (G^tV^{-1}G)^{-1}G^tV^{-1}(y-y_0)$$

Covariance of New Model Parameters p

Updated model parameter **p** for experimental points **y**~f(**p**)

$\mathbf{p}= \mathbf{p_0} + (\mathbf{G}^t\mathbf{V}^{-1}\mathbf{G})^{-1}\mathbf{G}^t\mathbf{V}^{-1}(\mathbf{y}-\mathbf{y_0})$

If we define $\mathbf{K}=(\mathbf{G}^t\mathbf{V}^{-1}\mathbf{G})^{-1}\mathbf{G}^t\mathbf{V}^{-1}$, $\mathbf{p}=\mathbf{p_0} + \mathbf{K}(\mathbf{y}-\mathbf{y_0})$ or $p_f=p_{0f}+\Sigma_{i=1,m}K_{fi}(y_i-y_{0i})$

For two linear combinations $f = \Sigma_{i=1,n} K_{fi}y_i$ and $g = \Sigma_{j=1,m} K_{gj}y_j$,

The covariance between f and g is

$V_{fg}= \Sigma_{i=1,n;\ j=1,m} K_{fi}K_{gj}V_{ij}$ (V_{ij}: covariance between y_i and y_j).

Therefore the covariance between p_f and p_g is

$M_{fg}= \Sigma_{i=1,n;\ j=1,m} K_{fi}K_{gi}V_{ij}$ or

$\mathbf{M}=\mathbf{KVK}^t =[(\mathbf{G}^t\mathbf{V}^{-1}\mathbf{G})^{-1}\mathbf{G}^t\mathbf{V}^{-1}]\mathbf{V}[(\mathbf{G}^t\mathbf{V}^{-1}\mathbf{G})^{-1}\mathbf{G}^t\mathbf{V}^{-1})^t]=(\mathbf{G}^t\mathbf{V}^{-1}\mathbf{G})^{-1}$

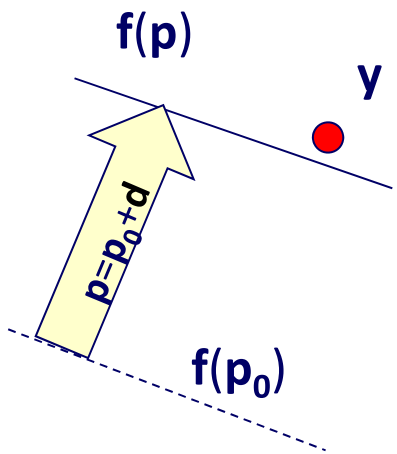
> $p = \Sigma_{i=1,n} a_i \cdot x_i \rightarrow V_{pq} \sim \Sigma_{i=1,n;\ j=1,m} g_i h_j V_{ij}$ $g_i=(\partial p/\partial x_i)_{xi=<xi>}$, $h_i=(\partial q/\partial y_j)_{yj=<yj>}$

Flow of Conventional Least Squares

Old model parameters
$\mathbf{p_0}=\{p_{0j}\}$

New experimental data
$\mathbf{y}=\{y_i\}$, $\mathbf{V}=\{\Delta y_i\}$

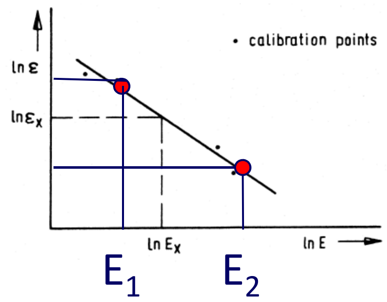
Minimum condition
$\chi^2=[\mathbf{y}-\mathbf{f}(\mathbf{p})]^t\mathbf{V}^{-1}[\mathbf{y}-\mathbf{f}(\mathbf{p})]=\min$



New model parameters and covariances
$\mathbf{p}=\{p_j\}=\mathbf{p_0}+(\mathbf{G}^t\mathbf{V}^{-1}\mathbf{G})^{-1}\mathbf{G}^t\mathbf{V}^{-1}(\mathbf{y}-\mathbf{y_0})$
$\mathbf{M}=(\mathbf{G}^t\mathbf{V}^{-1}\mathbf{G})^{-1}$

Adjustment Makes Correlation between Parameters!

Adjustment to experimental points **y** creates correlation **M** between model parameters **p**!



Example: Efficiency curve ε=f(p)

Model (fitting) parameters **p** were determined to reproduce efficiency values at several calibration points by standard sources.

Even if these calibration points are independent each other, **p** gets correlation **M**.

Efficiencies at two energies interpolated by the model (fitting) $\varepsilon_1=f(E_1,\mathbf{p})$ and $\varepsilon_2=f(E_2,\mathbf{p})$ are <u>also</u> correlated through **M**.

Weighted Average

Conventional least squares:

n experimental points $y_i \pm \Delta y_i$, $v_{ii}=\Delta y_i^2$, $(V_{ii})^{1/2}=\Delta y_i$ (i=1,n)

→ find p_j (j=1,m) which minimize $\Sigma_{i=1,n}[y_i-f_i(p_1,\ldots,p_m)]^2/V_{ii}$.

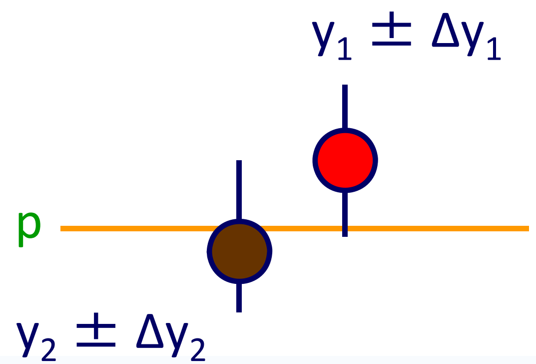
Weighted average p of experimental points y_1 and y_2:

2 experimental points $y_1 \pm \Delta y_1$ and $y_2 \pm \Delta y_2$, $v_{11}=\Delta y_1^2$ and $v_{22}=\Delta y_2^2$.

→ find p which minimize $Q=\chi^2=(y_1-p)^2/V_{11}+(y_2-p)^2/V_{22}$

→ $\partial Q/\partial p=0$

→ $p=[(y_1/V_{11})+(y_2/V_{22})]/[(1/V_{11})+(1/V_{22})]$



Weighted Average

From 2 experiment x-sections $y_1 \pm \Delta y_1$ and $y \pm \Delta y_2$.

$p=[(y_1/V_{11})+(y_2/V_{22})]/[(1/V_{11})+(1/V_{22})]$

$=[(y_1/\Delta y_1^2)+(y_2/\Delta y_2^2)]/[(1/\Delta y_1^2)+(1/\Delta y_2^2)]$

If we define $w_1=1/\Delta y_1^2$ and $w_2=1/\Delta y_2^2$, this equation becomes

$\mathbf{p=(w_1 y_1+w_2 y_2)/(w_1+w_2)}$

This is a linear combination of y_1 and y_2, and therefore

$(\Delta p)^2=M_{pp}=[w_1/(w_1+w_2)]\Delta y_1^2+[w_2/(w_1+w_2)]\Delta y_2^2$

$=\Delta y_1^2 \Delta y_2^2/(\Delta y_1^2+\Delta y_2^2)$

This is the **weighted average** of two experimental data points which have variances without correlation.

$p = \Sigma_{i=1,n}\, a_i \cdot x_i \rightarrow \langle p\rangle= \Sigma_{i=1,n}\, a_i\langle x_i\rangle,\ V_{pp}=\langle p^2\rangle-\langle p\rangle^2=\Sigma_{i=1,n} a_i^2 V_{ii} +2\Sigma_{i=1,n;\ j=1,n;i<j}\, a_i a_j V_{ij}$

Example of Weighted Average

Weighted average:

$y_1 \pm \Delta y_1$, $y_2 \pm \Delta y_2$

$\rightarrow$ $p=[(y_1/\Delta y_1^2)+(y_2/\Delta y_2^2)]/[(1/\Delta y_1^2)+(1/\Delta y_2^2)]$,

$(\Delta p)^2=\Delta y_1^2 \Delta y_2^2/(\Delta y_1^2+\Delta y_2^2)$

Example (Exercise):

Calculate the weighted average and its standard deviation for

$y_1 \pm \Delta y_1$ = **1.0±0.1** and $y_2 \pm \Delta y_2$ = **1.2±0.2**

Weighted average:

$p = [(1.0/0.1^2)+(1.2/0.2^2)]/[(1/0.1^2)+(1/0.2^2)]=$**1.04.**

Variance:

$(\Delta p)^2=(0.1^2 \cdot 0.2^2)/(0.1^2+0.2^2)=0.008$ $\Delta p=$**0.089**.

Off-Diagonal Weighted Average

Conventional least squares:

n experimental points $y_i \pm \Delta y_i$, $v_{ii}=\Delta y_i^2$, $(v_{ii})^{1/2}=\Delta y_i$ (i=1,n)

→ find p_j (j=1,m) minimizing $\Sigma_{i=1,n}[y_i-f_i(p_1,\ldots,p_m)]^2/V_{ii}$

or $[\mathbf{y}-\mathbf{f}(\mathbf{p})]^t\mathbf{V}^{-1}[\mathbf{y}-\mathbf{f}(\mathbf{p})]$ in vector-matrix notation.

This expression is valid also when **V** has off-diagonal elements.

Off-diagonal weighted average p of experimental points y_1 and y_2:

2 experimental points $y_1 \pm \Delta y_1$ and $y_2 \pm \Delta y_2$ with covariance **V**.

→ find p minimizing $Q=\chi^2=[\mathbf{y}-\mathbf{f}(p)]^t\mathbf{V}^{-1}[\mathbf{y}-\mathbf{f}(p)]=\Sigma_{i,j=1,2}(y_i-p)V^{-1}_{ij}(y_j-p)$

→ $\partial Q/\partial p=0$ → .... → $\mathbf{p= [(V_{22}-V_{12})y_1+(V_{11}-V_{12})y_2]/(V_{11}+V_{22}-2V_{12})}$

This is the **weighted average** of two experimental data points which have covariance (with correlation).

Variance of Off-Diagonal Weighted Average

Off-diagonal weighted average of y_1 and y_2:

$p= [(V_{22}-V_{12})y_1+(V_{11}-V_{12})y_2]/(V_{11}+V_{22}-2V_{12})$

If we set $a= (V_{22}-V_{12})/(V_{11}+V_{22}-2V_{12})$ and $b=(V_{11}-V_{12})/(V_{11}+V_{22}-2V_{12})$,

$p= (ay_1+by_2)$ – linear combination of y_1 and y_2.

Applying the error propagation law for linear combination,

$(\Delta p)^2=v_{pp}=a^2V_{11}+b^2V_{22}+ 2abV_{12}$

$= [(V_{22}-V_{12})^2V_{11}+(V_{11}-V_{12})^2V_{22}+2(V_{22}-V_{12})(V_{11}-V_{12})V_{12}]$
$/(V_{11}+V_{22}-2V_{12})^2.$

$= [V_{11}V_{22}-(V_{12})^2]/(V_{11}+V_{22}-2V_{12})$

$p = \Sigma_{i=1,n}\, a_i \cdot x_i \rightarrow \langle p\rangle= \Sigma_{i=1,n}\, a_i\langle x_i\rangle,\ V_{pp}=\langle p^2\rangle-\langle p\rangle^2=\Sigma_{i=1,n}a_i^2V_{ii} +2\Sigma_{i=1,n;\ j=1,n;i<j}\, a_ia_jV_{ij}$

Diagonal and Off-Diagonal Weighted Average

Off-diagonal weighted average

$$p= [(V_{22}-V_{12})y_1+(V_{11}-V_{12})y_2]/(V_{11}+V_{22}-2V_{12})$$

$$(\Delta p)^2=[V_{11}V_{22}-(V_{12})^2]/(V_{11}+V_{22}-2V_{12})$$

If we set $\mathbf{V_{12}=0}$, then we obtain

$$p= [V_{22}y_1+V_{11}y_2]/(V_{11}+V_{22})$$

$$(\Delta p)^2=(V_{11}V_{22})/(V_{11}+V_{22}).$$

This is the (diagonal) weighted average and its variance (as should be!).

Off-Diagonal Weighted Average: Case Studies

$p= [(V_{22}-V_{12})y_1+(V_{11}-V_{12})y_2]/(V_{11}+V_{22}-2V_{12})$

Case 1. Average of the <u>same experimental data point</u>

$y_1=y_2$

$\Delta y_1=\Delta y_2$ (i.e., $V_{11}=V_{22}$)

$c_{12}=V_{12}/(V_{11}V_{22})^{1/2}=1$ (fully correlated – same experiment)

$\rightarrow V_{11}=V_{22}=V_{12}$

This makes the covariance matrix V <u>singular</u> ($|V|=V_{11}+V_{22}-2V_{12}=0$).

Off-Diagonal Weighted Average: Case Studies

$p= [(V_{22}-V_{12})y_1+(V_{11}-V_{12})y_2]/(V_{11}+V_{22}-2V_{12})$

$(\Delta p)^2=[V_{11}V_{22}-(V_{12})^2]/(V_{11}+V_{22}-2V_{12})$

Case 2. Average of two experimental points with <u>same variance</u>

$\Delta y_1=\Delta y_2$ (i.e., $V_{11}=V_{22}$)

$p= [(V_{11}-V_{12})y_1+(V_{11}-V_{12})y_2]/(2V_{11}-2V_{12})= (y_1+y_2)/2$

$(\Delta p)^2=M_{pp}=[V_{11}{}^2-(V_{12})^2]/(2V_{11}-2V_{12})=(V_{11}+V_{12})/2$

Off-Diagonal Weighted Average: Extreme Cases

Case 2. Average of two experimental points with same variance

$p= [(V_{11}-V_{12})y_1+(V_{11}-V_{12})y_2]/(2V_{11}-2V_{12})= (y_1+y_2)/2$

$(\Delta p)^2=[V_{11}^2-(V_{12})^2]/(2V_{11}-2V_{12})=(V_{11}+\mathbf{V_{12}})/2$

W. Mannhart (WM11 Appendix 2):

For uncorrelated data (V_{12}=0) the variance of the average (Δp) is give by half of the common variance of both data.

For correlated data ($V_{12}\neq0$) the variance of the average (Δp) remains above this value for a positive value of V_{12}.

This means a positive correlation of data lessens the uncertainty reduction of an evaluated result.

In the past this effect was sometimes responsible for inconsistencies in evaluated data, since the neglect of covariances gave too strong a reduction of the evaluated final uncertainties.

Off-Diagonal Weighted Average: Case Studies

$p= [(V_{22}-V_{12})y_1+(V_{11}-V_{12})y_2]/(V_{11}+V_{22}-2V_{12})$

Case 3. $V_{11}<V_{12}<V_{22}$ (or $V_{22}<V_{12}<V_{11}$)

The average p is larger or smaller than BOTH y_1 and y_2(!).

Proof:

If we define $a=V_{22}-V_{12}$ (>0) and $b=V_{12}-V_{11}$ (>0), $p=(ay_1-by_2)/(a-b)$.

$(p-y_1)(p-y_2)=[(ay_1-by_2)/(a-b)-y_1]\cdot[(ay_1-by_2)/(a-b)-y_2]$

$= ab(y_1-y_2)^2/(a-b)$

>0

Therefore both $p - y_1$ and $p - y_2$ are positive or negative.

Generalized Least Squares

In conventional least squares, we assumed that

- covariance matrix **V** for experimental data **y** is diagonal.
- no variance-covariance for old model parameters **p**.

→

Generalized least squares approach considers

- correlation in experimental data (i.e. off-diagonal elements of **V**)
- covariance of old fitting parameters **p**

Let us see the answer of this problem *without proves*!

Conventional v.s. General: Minimum Conditions

Conventional least squares

$$\chi^2=[\mathbf{y}-\mathbf{f}(\mathbf{p})]^t\mathbf{V}^{-1}[\mathbf{y}-\mathbf{f}(\mathbf{p})]=\min$$

New model parameter **p** should be consistent with experimental data **y** through the model **f**

Generalized least squares

$$\chi^2=[\mathbf{y}-\mathbf{f}(\mathbf{p})]^t\mathbf{V}^{-1}[\mathbf{y}-\mathbf{f}(\mathbf{p})]+[\mathbf{p}-\mathbf{p}_0]^t\mathbf{M}_0^{-1}[\mathbf{p}-\mathbf{p}_0]=\min$$

New model parameters **p** should be consistent with experimental data **y**, through the model **f**, *and* should be consistent with old model parameters $\mathbf{p}_0$.

Formulae of Generalized Least-Squares

- **Problem:**

  Update model parameters $\mathbf{p_0}=\{p_{0j}\}$ (covariance $\mathbf{M_0}$) to $\mathbf{p}=\{p_j\}$ by experimental points $\mathbf{y}=\{y_i\}$ (covariances $\mathbf{V}$) (i=1,n).

- Model $\mathbf{y}$~f($\mathbf{p}$) can be expanded around prior model parameters $\mathbf{p_0}$ as

  $f(\mathbf{p}) \sim f(\mathbf{p_0})+\mathbf{G}(\mathbf{y}-\mathbf{y_0})$, $\mathbf{G}= \partial f(\mathbf{p})/\partial \mathbf{p}_{|\mathbf{p}=\mathbf{p0}}=\mathbf{G}$.

- Updated model parameters **p** and their covariance **M** are

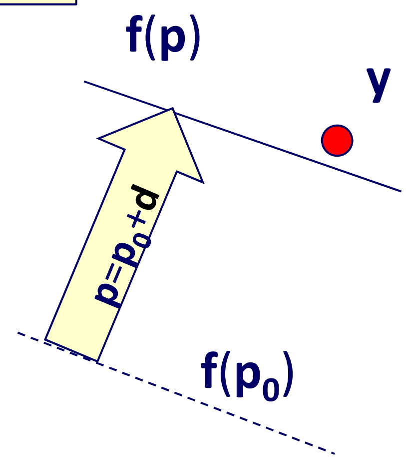
  $\mathbf{p}= \mathbf{p_0}+\mathbf{M_0}\mathbf{G}^t(\mathbf{G}\mathbf{M_0}\mathbf{G}^T+\mathbf{V})^{-1}(\mathbf{y}-\mathbf{y_0})$

  $\mathbf{M}=\mathbf{M_0}-\mathbf{M_0}\mathbf{G}^t(\mathbf{G}\mathbf{M_0}\mathbf{G}^t+\mathbf{V})^{-1}\mathbf{G}\mathbf{M_0}$

Flow of Generalized Least-Squares

Old model parameters
$\mathbf{p_0}=\{p_{0j}\}$
with covariance $\mathbf{M_0}$

New experimental data
$\mathbf{y}=\{y_i\}$
with covariance $\mathbf{V}$



Minimum condition

$$\chi^2=[\mathbf{y}-f(\mathbf{p})]^t\mathbf{V}^{-1}[\mathbf{y}-f(\mathbf{p})]+[\mathbf{p}-\mathbf{p_0}]^t\mathbf{M_0}^{-1}[\mathbf{p}-\mathbf{p_0}]=\min$$

New fitting parameter and covariances

$$\mathbf{p}=\{p_j\}=\mathbf{p_0}+\mathbf{M_0G}^t(\mathbf{GM_0G}^t+\mathbf{V})^{-1}(\mathbf{y}-\mathbf{y_0})$$

$$\mathbf{M}=\mathbf{M_0}-\mathbf{M_0G}^t(\mathbf{GM_0G}^t+\mathbf{V})^{-1}\mathbf{GM_0}$$

Summary

Generalized least-squares:

Prior model parameters $\mathbf{p_0}$ (covariance $\mathbf{M_0}$) are updated by experimental data $\mathbf{y}$ (covariance $\mathbf{V}$) by

$$\mathbf{p}=\mathbf{p_0}+\mathbf{M_0G^t(GM_0G^t+V)^{-1}(y-y_0)}$$

$$\mathbf{M}=\mathbf{M_0}-\mathbf{M_0G^t(GM_0G^t+V)^{-1}GM_0}$$

Weighted average:

The weighted average of two data y_1 and y_2 and its variance are

$$p= [(V_{22}-V_{12})y_1+(V_{11}-V_{12})y_2]/(V_{11}+V_{22}-2V_{12})$$

$$(\Delta p)^2=[V_{11}V_{22}-(V_{12})^2]/(V_{11}+V_{22}-2V_{12})$$

The off-diagonal element V_{12} (i.e., correlation) plays an important role.

Exercise (~13:00)

Choose the following tasks as you like:

- Paper 2.2 (Weighted average with and without covariance)
- Paper 2.3 (Preparation of input for least-squares analysis)
- Paper 2.4 (Weighted average with covariance).
- Go for lunch.