Notes on least squire method

Assumptions: you have n measurements y_i with known errors σ_i $(i=1,2,\ldots,n)$ and a model so that you can predict what should be true values for these measurements $f_i(a_{1,} a_{2,} \ldots, a_m)$. Usually the model is not fixed but has a few unknown parameters designated here as a_j, $j=1,2,\ldots,m$.

The least squire method allow you to get the "best" values of the parameters in a sense that with these values your model describe your measurements in the best way. And the best way is understood as all f_i are as close as possible to the corresponding y_i. Or, mathematically, the following expression should be minimum:

$$\chi^2=\sum_{i=1}^{n}\left(\frac{y_i-f_i(a_{1,}\ldots,a_m)}{\sigma_i}\right)^2\to min$$

Recipe

In the simple case when f_i depend on a_j linearly

$$f_i=h_{i1}a_1+h_{i2}a_2+\ldots+h_{im}a_m$$

the following recipe will solve the problem:

1. compose a following matrix $\boldsymbol{A}$ and a vector $\boldsymbol{Y}$

   $$A_{ij}=\sum_{k=1}^{n} g_k h_{ki} h_{kj} \quad \text{and} \quad Y_i=\sum_{k=1}^{n} g_k h_{ki} y_k, \quad \text{where} \quad g_k=1/\sigma_k^2$$

2. the best values of $\boldsymbol{a}=\|a_{1,}\ldots,a_m\|^T$ are obtain from the following linear equation: $\boldsymbol{A}\boldsymbol{a}=\boldsymbol{Y}$
3. $\boldsymbol{A}^{-1}$ is also a covariance matrix of the estimated parameters $\boldsymbol{a}$.

Example

Four detectors located on a horizontal plane are used to measure the time t_i of arrival of a cosmic ray shower front. The coordinate system on the plane is chosen so that one of the detector is in the center of the system, others have coordinates $(x_{2,}y_2),(x_{3,}y_3),(x_{4,}y_4)$. Time is measured in respect to the (0,0) detector, so t_1 is always 0. The model here is that the shower front is a plane which can move at any angle toward the detector's plane. In the chosen coordinate system any point *(x,y,z)* of this shower plane at a moment *t* should satisfy the equation:

$n_x x+n_y y+n_z z=-v\,t$, where $n_{x,y,z}$ are projections of a unit vector perpendicular to the shower plane, v is the velocity it moves with.

In this example measurements are $y_i = t_i$, the predicted times, $f_i(n_x, n_y) = -\frac{x_i}{v} n_x - \frac{y_i}{v} n_y$ (z=0 for all detectors), depend on parameters (n_x, n_y) to be determined, $i = 1,2,3$. So $h_{i1} = -\frac{x_i}{v}$, $h_{i2} = -\frac{y_i}{v}$. According to the recipe let's make a system of linear equations:

$$\begin{Vmatrix} \sum_{k=1}^{3} g_k \left(\frac{x_k}{v}\right)^2 & \sum_{k=1}^{3} g_k \frac{x_k y_k}{v^2} \\ \sum_{k=1}^{3} g_k \frac{x_k y_k}{v^2} & \sum_{k=1}^{3} g_k \left(\frac{y_k}{v}\right)^2 \end{Vmatrix} \begin{Vmatrix} n_x \\ n_y \end{Vmatrix} = \begin{Vmatrix} \sum_{k=1}^{3} g_k \frac{x_k}{v} t_k \\ \sum_{k=1}^{3} g_k \frac{y_k}{v} t_k \end{Vmatrix}$$

Since we consider σ_k to be the same for all measurements $\sigma_k = \sigma \approx 2.5 ns$, $g_k = g = 1/\sigma^2$ can be taken out from both sides of the equation and cancels. Also *v* can be taken out:

$$\frac{1}{v} \begin{Vmatrix} \sum_{k=1}^{3} x_k^2 & \sum_{k=1}^{3} x_k y_k \\ \sum_{k=1}^{3} x_k y_k & \sum_{k=1}^{3} y_k^2 \end{Vmatrix} \begin{Vmatrix} n_x \\ n_y \end{Vmatrix} = \begin{Vmatrix} \sum_{k=1}^{3} x_k t_k \\ \sum_{k=1}^{3} y_k t_k \end{Vmatrix}$$

This is a system of the type:

$\begin{cases} A n_x + B n_y = E \\ B n_x + C n_y = F \end{cases}$, with all coefficients known...

Unfinished section

This section is unfinished. Don't bother to read...

$$(\boldsymbol{f}(\boldsymbol{a}) - \boldsymbol{y})^{\boldsymbol{T}} \boldsymbol{G} (\boldsymbol{f}(\boldsymbol{a}) - \boldsymbol{y}) \rightarrow min$$

$$\boldsymbol{f}(\boldsymbol{a}) = \boldsymbol{H}(\boldsymbol{a} - \boldsymbol{a_0}) + \boldsymbol{f}(\boldsymbol{a_0}) \qquad H_{ij} = \frac{\partial f_i}{\partial a_j} \qquad \boldsymbol{\Theta = a - a0} \qquad \boldsymbol{-y_0 = f(a_0) - y}$$

$$(\boldsymbol{H\Theta} - \boldsymbol{y_0})^{\boldsymbol{T}} \boldsymbol{G} (\boldsymbol{H\Theta} - \boldsymbol{y_0}) \rightarrow min$$

$$\boldsymbol{2 H^T G (H\Theta - y_0) = 0}$$

$$\boldsymbol{(H^T G H)\Theta = H^T G y_0}$$

$$\boldsymbol{\Theta = (H^T G H)^{-1} H^T G y_0}$$

$$\boldsymbol{V}_{\Theta}=(\boldsymbol{H}^{T}\boldsymbol{G}\boldsymbol{H})^{-1}$$

$$A_{ij}=\sum_{k=1}^{n}\sigma_k\frac{\partial f_k}{\partial a_i}\frac{\partial f_k}{\partial a_j} \quad \text{and} \quad Y_i=\sum_{k=1}^{n}\sigma_k\frac{\partial f_k}{\partial a_j}y_k$$

$$A=\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1\mathrm{m}} \\ a_{21} & a_{22} & \dots & a_{2\mathrm{m}} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{Vmatrix}$$