Application of a generalized Method of Least
Squares for Kinematical Analysis of
Tracks in Bubble Chambers

by

R. Böck
Propriété littéraire et scientifique réservée pour tous les pays du monde. Ce document ne peut être reproduit ou traduit en tout ou en partie sans l’autorisation écrite du Directeur général du CERN, titulaire du droit d’auteur. Dans les cas appropriés, et s’il s’agit d’utiliser le document à des fins non commerciales, cette autorisation sera volontiers accordée.

Le CERN ne revendique pas la propriété des inventions brevetables et dessins ou modèles susceptibles de dépôt qui pourraient être décrits dans le présent document; ceux-ci peuvent être librement utilisés par les instituts de recherche, les industriels et autres intéressés. Cependant, le CERN se réserve le droit de s’opposer à toute revendication qu’un usager pourrait faire de la propriété scientifique ou industrielle de toute invention et tout dessin ou modèle décrits dans le présent document.

© Copyright CERN, Genève, 1960

Literary and scientific copyrights reserved in all countries of the world. This report, or any part of it, may not be reprinted or translated without written permission of the copyright holder, the Director-General of CERN. However, permission will be freely granted for appropriate non-commercial use. If any patentable invention or registrable design is described in the report, CERN makes no claim to property rights in it but offers it for the free use of research institutions, manufacturers and others. CERN, however, may oppose any attempt by a user to claim any proprietary or patent rights in such inventions or designs as may be described in the present document.
Application of a generalized Method of Least Squares for Kinematical Analysis of Tracks in Bubble Chambers

by

R. Böck
Application of a generalized Method of Least
Squares for Kinematical Analysis of
Tracks in Bubble Chambers

by
R. Böck

I.

In applying least squares methods to fit the geometrically reconstructed tracks of a bubble chamber event into a kinematically reasonable hypothesis, i.e. to force the momentum four-vector-conservation for the point of interaction, difficulties arose in that some of the entry quantities defining the tracks were unmeasured or poorly defined. Generally one used one or more of the four constraint equations to determine the unknown variable(s) and modified the remaining equations accordingly. See the publications of the Alvarez Group, Berkeley, on this subject: UCRL 9097 (J.P. Berge, P.T. Solmitz, H.D. Taft) and MEMO 86 (J.P. Berge), dealing with the IHEP 704 - programme "GUTS."

It is possible to treat the problem in a more general form so that a computer programme will require only a very simple entry parameter to define which of the variables are measured and which have, hence, to enter the minimum function. This most general case will be treated in section V of this note. To get familiar with the matrix notation used and for completeness I have also described the simpler cases of fitting functions of unknowns to a series of single measurements (III) and of fitting

9170/p
measurements into constraints with no unknown variables (IV).
I wish to point out also that section V is but a known procedure.
The idea of this note is to present a condensed extract of what
can be found in various publications on statistics or least
squares methods.

II.

Definitions
Throughout this note the following notation is
used:

\[ m = (a, 1) \text{ matrix (a-vector) of measured variables} \]
\[ c = \text{a-vector of corrections to } m, \text{ so that } m + c \text{ satisfies} \]
the conditions required. The elements of \( c \) are assumed
to be normally (Gaussian) distributed, but not necessarily
without correlation.

\[ G_m^{-1} = (a, a) \text{ error matrix, where } (G_m^{-1})_{ii} \text{ (i-th element of the} \]
\[ \text{diagonal) stands for the variance of the i-th element of} \]
\[ m \text{ and } (G_m^{-1})_{ij} \text{ for the covariance of i-th and j-th element} \]
of \( m \). Generally \( G_k^{-1} \) stands for the error matrix of \( k \).
The inverse of an error matrix is called a weighting matrix.
Consider that by definition \( G_k^{-1} \) is symmetric.

\[ x = b \text{-vector of unknowns} \]
\[ \bar{x} = b \text{-vector of approximations for } x \]
\[ \Delta x = b \text{-vector of corrections for } \bar{x} \]
\[ f(m,x) = c \text{-vector of functions of } m \text{ and } x \]

\[ \frac{\partial f}{\partial m} = (c, a) \text{ matrix of partial derivatives, defined by } \frac{\partial f}{\partial m} = \]
\[ \frac{\partial f_i}{\partial m_j} \]

\( M = \text{function to be minimized} \)

9170/p
\( \alpha = \) c-vector of Lagrangean multipliers
\( \mathbf{I} = \) unit matrix of required dimensions

superscripts: \( \nu \) stands for iteration number
\( \mathbf{T} \) " matrix transpose
\( -1 \) " inversion

III.

The first problem to be considered is to fit a number of functions of unknowns to a series of measurements.

\[
f(x) = m + c \quad \text{(III.1)}
\]

We have the additional requirement

\[
M = \mathbf{c}^T \mathbf{G}_m \mathbf{c} = \text{minimum} \quad \text{(III.1a)}
\]

Assume we have approximations for the unknowns, so that we can expand

\[
f(\bar{x}) + \frac{\partial f}{\partial \bar{x}} \Delta x = m + c \quad \text{(III.2a)}
\]

we introduce \( A = \frac{\partial f}{\partial x} \) and \( r = f(\bar{x}) - m \) (residuals)

to get

\[
A \Delta x + r = c \quad \text{(III.2)}
\]

Introducing (2) into the minimum function we get

\[
M = (A \Delta x + r)^T \mathbf{G}_m (A \Delta x + r) \quad \text{(III.3)}
\]

Our requirement is to choose the \( \Delta x \) such that \( M \) is a minimum, i.e. \( \frac{\partial M}{\partial \Delta x} = 0 \) (b-zero vector). By definition of the differentiation symbol one obtains

\[
\frac{\partial M}{\partial \Delta x} = 2(A \Delta x + r)^T \mathbf{G}_m A = 0 \quad \text{(III.4)}
\]
which gives the solution

$$\Delta x = -(A^T G_m A)^{-1} A^T G_m r$$  \hspace{1cm} (III.5)$$

Using (III.2) one has the solution for \( c \). (III.1a) and (III.3) give an identity. After introducing (III.5) into (III.3) one can write

$$M = c^T G_m c = r^T G_m r + r^T G_m A \Delta x$$  \hspace{1cm} (III.6)$$

which is useful as a numerical check.

If \( f \) is linear, the expansion (III.2) is valid for any \( \bar{x} \) and one step will give final answers. If \( f \) contains one or more non-linear functions, the procedure is iterative and one obtains

$$\bar{x}^{\nu+1} = \bar{x}^{\nu} + \Delta x^{\nu+1}.$$  \hspace{1cm} (III.7)$$

After a suitable exit criterion has been satisfied, one obtains the error matrices for the fitted quantities by

$$G_x^{-1} = G_x^{-1} \frac{\partial \Delta x}{\partial m} = (\frac{\partial \Delta x}{\partial m})^T G_m^{-1} \frac{\partial \Delta x}{\partial m}$$

which is just the law of error propagation. Hence

$$G_x^{-1} = \left[(A^T G_m A)^{-1} A^T G_m \right] G_m^{-1} \left[(A^T G_m A)^{-1} A^T G_m \right]^T$$

$$G_x^{-1} = (A^T G_m A)^{-1}$$  \hspace{1cm} (III.7)$$

and accordingly with \( \frac{\partial(m + c)}{\partial m} = A \frac{\partial \Delta x}{\partial m} = A(A^T G_m A)^{-1} A^T G_m \)

$$G_{(m + c)}^{-1} = A(A^T G_m A)^{-1} A^T$$  \hspace{1cm} (III.8)$$

The number of degrees of freedom is \( a - b \).

IV.

Next we come to the procedure to fit a number of completely measured variables into a scheme of constraints without determination of one or more unmeasured quantities. Our starting equations now are

9170/p
\[ f(m + c) = 0 \quad (IV.1) \]

with the minimum condition
\[ M = c^T G_m c = \text{minimum} \quad (IV.1a) \]

we assume approximations for \( m + c \), namely \( m + \overline{c} \), which is necessary, if \( f \) is non-linear and the \( c \) are not negligible) and expand
\[ f(m + \overline{c}) + \frac{\partial f}{\partial (m + \overline{c})} (c - \overline{c}) = 0 \quad (IV.2a) \]

with \( B = \frac{\partial f}{\partial (m + \overline{c})} \) and \( r = f(m + \overline{c}) - B\overline{c} \)

we have
\[ Bc + r = 0 \quad (IV.2) \]

The usual procedure is now to introduce the Lagrangeian multipliers in \((IV.1a)\) and write
\[ M = c^T G_m c + 2\alpha^T (Bc + r) \quad (IV.3) \]

The requirement \( \frac{\partial M}{\partial \alpha} = 0 \) leads us to the expression
\[ c = - G_m^{-1} B^T \alpha \quad (IV.4) \]

which by introducing into \((IV.2)\) gives the solution
\[ \alpha = (B G_m^{-1} B^T)^{-1} r \quad (IV.5) \]

The \( M \)-identity now is reduced to
\[ M = c^T G_m c = \alpha^T \quad (IV.6) \]

The non-linearity in \( f \) requires iterative steps again and each step yields a \( \overline{c} P + 1 = - G_m^{-1} B^T \alpha \). \( \overline{c} = c \) is, of course, already a good approximation. The error matrix for the fitted quantities \( m + c \) is given by using the relation
\[ \frac{\partial (m + c)}{\partial m} = I - G_m^{-1} B^T \frac{\partial \alpha}{\partial m} \]
\[ = I - G_m^{-1} B^T (B G_m^{-1} B^T)^{-1} B \]

9170/p
(Since \( \frac{\partial x}{\partial m} = B \)) so that

\[
G^{-1}_{(m + c)} = G^{-1}_m - G^{-1}_m B^T (B G^{-1}_m B^T)^{-1} B G^{-1}_m \quad \text{(IV.7)}
\]

Note that \( G^{-1}_{(m + c)} \) is made up of \( G^{-1}_m - G^{-1}_c \), hence

\[
(G^{-1}_{(m + c)})_{ii} \leq (G^{-1}_m)_{ii}
\]

The number of degrees of freedom is \( c \).

V.

The extension to the most general case is now obvious. We assume a function of both the measured (fitted) and the unknown variables.

\[
f(x, m + c) = 0 \quad \text{(V.1)}
\]

For both types of variables approximations are known and we can again expand as in III and IV:

\[
f(\bar{x}, m + \bar{c}) + \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial (m + \bar{c})} (c - \bar{c}) = 0 \quad \text{(V.2a)}
\]

which leads to our starting equations

\[
A \Delta x + B c + r = 0 \quad \text{(V.2)}
\]

with \( r = f(\bar{x}, m + \bar{c}) - B \bar{c} \), and \( A \) and \( B \) as in III and IV.

In addition we have the minimum condition

\[
M = c^T G_m c = \text{minimum} \quad \text{(V.1a)}
\]

Assuming first the \( \Delta x \) as known we get the case of section IV and using the formulae (IV.2) to (IV.5) we obtain

\[
M = c^T G_m c + 2 \alpha^T (A \Delta x + B c + r) \quad \text{(V.3)}
\]

\[
= c^T G_m c + 2 \alpha^T (B c + r')
\]

9170/p
\[ c = - G_m^{-1} B^T \alpha \]  \hspace{1cm} (V.4)

and
\[ \alpha = (B G_m^{-1} B^T)^{-1} r = (B G_m^{-1} B^T)^{-1} (A \Delta x + r) \]  \hspace{1cm} (V.5)

We define for convenience \( G_B^{-1} = B G_m^{-1} B^T \). \( G_B^{-1} \) is the error matrix for the elements of \( Bm \). Introducing our solution for \( \alpha \) (V.5) into (V.3) we get
\[ M = (A \Delta x + r)^T G_B (A \Delta x + r) \]  \hspace{1cm} (V.6)

and from now on we follow section III, equations (III.3) to (III.5) to obtain
\[ \frac{\partial M}{\partial \Delta x} = 2 (A \Delta x + r)^T G_B A = 0 \]  \hspace{1cm} (V.7)

and
\[ \Delta x = - (A^T G_B A)^{-1} A^T G_B r \]  \hspace{1cm} (V.8)

Thus the whole problem is solved. One first has to calculate both \( A \) and \( B \) and by some matrix multiplications the matrices \( G_m^{-1} B, G_B^{-1}, G_B A \) and \( A^T G_B A \). Starting at (V.8) one obtains first the \( \Delta x \) and then from (V.5) the \( \alpha \), and from (V.4) the \( c \).

The \( M \)-identity takes the form
\[ M = c^T G_m c = \alpha^T (A \Delta x + r) \]  \hspace{1cm} (V.9)

The requirements in the case of non-linear constraint equations are the same as in the previous sections.

Deriving finally the error matrices for the fitted \( x \) and \( m + c \) we obtain:
\[ \frac{\partial \Delta x}{\partial m} = - (A^T G_B A)^{-1} A^T G_B B \]

hence
\[ G_x^{-1} = G_m^{-1} \frac{\partial \Delta x}{\partial m} = (A^T G_B A)^{-1} \]  \hspace{1cm} (V.10)

9170/p
\[
\frac{\partial (m + c)}{\partial m} = I - G^{-1}_m B^T \frac{\partial g}{\partial m} \\
= I - G^{-1}_m B^T \left\{ G_B A \frac{\partial A}{\partial m} + G_B B \right\} \\
= I - G^{-1}_m B^T \left\{ -A(A^T G_B A)^{-1} A^T G_B B + B \right\}
\]

hence

\[
G^{-1}_B (m + c) = G^{-1}_m - G^{-1}_m B^T G_B B G^{-1}_m + G^{-1}_m B^T G_B A (A^T G_B A)^{-1} A^T G_B B G^{-1}_m 
\]

(V.11)

\[
= G^{-1}_m - G^{-1}_c
\]

Sometimes the error matrix for all adjusted quantities is required. The matrix of correlations between \((m + c)\) and \(x\) must then be calculated

\[
C(m + c), x = \left( \frac{\partial (m + c)}{\partial m} \right) G^{-1}_m \left( \frac{\partial x}{\partial m} \right)^T \\
= -G^{-1}_m B^T G_B A (A^T G_B A)^{-1}
\]

(V.12)

\(G^{-1}, G(m + c)\) and \(C(m + c), x\) (and \(C^T(m + c), x = C_x, (m + c)\)) together form the final error matrix. The number of degrees of freedom is \(c - b\). As can be seen, \(a\) is not the essential number for the determination of redundant information. There is, of course, the restriction that at least \(c\) out of the \(a\) measured quantities need to be linearly independent with respect to the constraints, to obtain a non-singular \(G^{-1}_B\).

VI.

Sections III and IV turn out to be just special cases of V. The case of section III is obtained by putting \(B = I\), and in the case of Section IV \(x\) is a \((0, 1)\) type vector respectively. All formulae derived in V become identical to those in III and IV under these assumptions.

9170/p
VII.

I come back to the application of the least squares method to the analysis of bubble chamber photographs, in particular to the fitting of geometrically reconstructed chamber tracks to the law of momentum and energy conservation at an interaction vertex. There are four constraints representing the conservation of the components of the momentum in a spatial cartesian coordinate system and the conservation of the total energy respectively. (The constraints, of course, can be formulated differently provided they are equivalent to the ones I mentioned). Each track is defined by 3 parameters, e.g. the momentum, the dip angle and the azimuthal angle. The choice of these parameters, however, will affect the results if the differences between measured and fitted quantities are not infinitely small. So in choosing them consideration should be given to how the tracks were obtained rather than to simplification of the analytical formulae.

In certain cases some of these parameters are not measurable and yet one wishes to fit the data on hand with a reduced number of degrees of freedom. Section V provides a means of treating such a case in a general way so that by a simple specification for each parameter (measured or unmeasured, single bit will do) the desired set of formulae will be applied. It is necessary to specify starting values for the unmeasured quantities, which, if badly determined, possibly can (other than in case of elimination) affect the convergence. But there is no need to branch inside the fit programme according to which quantities are measured or not. (Practically there is usually a necessity to branch, namely if the matrix routines used do not treat zero-columns properly or do not invert a $1 \times 1$ matrix).
The residuals and the matrix of derivatives can be formed as if one had to do with section III; only then the specifications for the parameters are tested and the derivatives are split up into those with respect to unmeasured quantities (matrix A in our notation) and the ones with respect to the measured ones (matrix B). Simultaneously one obtains, of course, the number of columns of each matrix which in our case can vary from 0 to 4 for A and from \((N - 4)\) to \(N\) for B \((N = \text{number of parameters involved} = 3 \times \text{number of tracks})\).

Apart from some data transfers this general l.s.f. method and the elimination method are identical in case of no unmeasured variables, so that there is only a negligible difference in computing time. With increasing number of unmeasured quantities one gains some time in case of elimination, as less equations have to be solved. Based on experience with an auto-code programme for the Ferranti Mercury computer one gains, expectedly, less time by the method outlined than by the elimination, but the crucial point is one still gains. The reduction in programming effort is considerable.
Literature:


F.T. Solmitz, J.P. Berge, H.D. Taft: Digital Analysis of Data from Bubble Chambers III. The Kinematical Analysis of Interaction Vertices, University of California, Lawrence Radiation Laboratory UCRL 9097 (1960)

J.P. Berge, IBM Program GUTS, University of California, Lawrence Radiation Laboratory, Physics Notes MEMO Nr. 86 (1959)