Fitting a Straight Line in Three-Dimensional Space by Total Least-Squares Adjustment

Kyle Snow and Burkhard Schaffrin

Division of Geodetic Science, School of Earth Sciences,
The Ohio State University, Columbus/Ohio, USA

Presented at the SIAM Annual Meeting, Chicago, 7–11 July 2014

- The typical parametric form of a 3-D line – six point and orientation parameters and parametric term t
- Roberts' (1988) representation of a 3-D line – four point and orientation parameters and parametric term t
- Using a Gauss-Markov Model with redundant data ($n/3$ points)
  - $6 + n/3$ parameters and 2 constraints
  - $4 + n/3$ parameters with Roberts' representation
- Using a Gauss-Helmert Model after eliminating $n/3$ parameters
  - Model definition and setup
  - A minimal parameterization for 3-D line fitting
- Total least-squares adjustment within the Gauss-Helmert Model
  - Numerical example – the "Flight of the Bumblebee"
- Conclusions and outlook
- References

- Line B in parametric form with parameter t can be written as

$$B = \{\boldsymbol{p} \,|\, \boldsymbol{p} = \boldsymbol{a} + t\boldsymbol{b},\ \boldsymbol{p} = [p_x, p_y, p_z]^T,\ -\infty < t < \infty\},$$

  with six point and orientation parameters:
  - $\boldsymbol{a} = [a_x,\ a_y,\ a_z]^T$ is a point on the line;
  - $\boldsymbol{b} = [b_x,\ b_y,\ b_z]^T$ defines the orientation of the line.
- But only four of the six parameters are independent!
- Roberts' conditions: Force uniqueness by requiring $\boldsymbol{b}$ to be a unit vector with $b_z \geq 0$, and force uniqueness on $\boldsymbol{a}$ by requiring it to be the point of nearest approach of line $\boldsymbol{B}$ to the origin.

$$\boldsymbol{b}^T\boldsymbol{b} = b_x^2 + b_y^2 + b_z^2 = 1, \quad b_z \geq 0 \Rightarrow b_z := \sqrt[+]{1 - b_x^2 - b_y^2}$$

$$\boldsymbol{a}^T\boldsymbol{b} = a_x b_x + a_y b_y + a_z b_z = 0$$

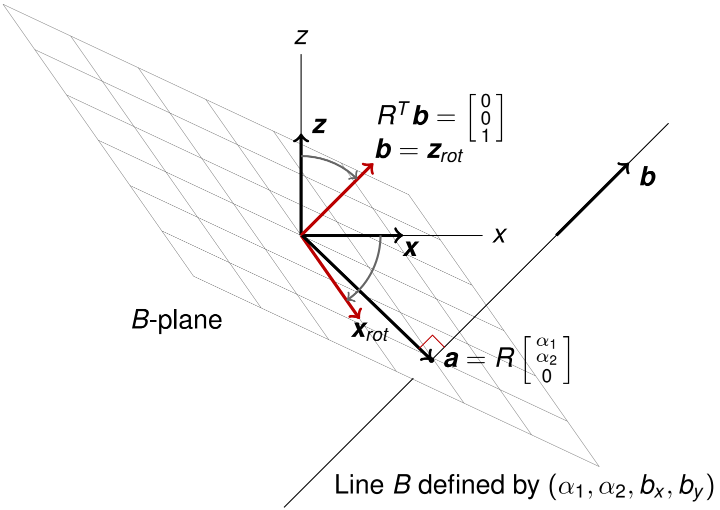
- Rotate the coordinate axes such that the positive z-axis is parallel with the *(unit)* orientation vector $\boldsymbol{b}$ and

$$\boldsymbol{a} = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = R\boldsymbol{\alpha} =: \underbrace{\begin{bmatrix} 1 - \frac{b_x^2}{1+b_z} & -\frac{b_x b_y}{1+b_z} & b_x \\ -\frac{b_x b_y}{1+b_z} & 1 - \frac{b_y^2}{1+b_z} & b_y \\ -b_x & -b_y & b_z \end{bmatrix}}_{\text{orthonormal}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ 0 \end{bmatrix} \Rightarrow \boldsymbol{\alpha} = R^T \boldsymbol{a}.$$

- The rotated x- and y-axis then span the so-called B-plane, which is normal to $\boldsymbol{b}$.
- The point $\boldsymbol{a}$ lies at the intersection of the line B and the B-plane and has coordinates $(\alpha_1, \alpha_2, 0)$ in the rotated coordinate system.
- Then, the point-and-orientation parameters for line B can be reduced to the four independent terms α_1, α_2, b_x, b_y with the (unit) orientation vector

$$\boldsymbol{b} = \left[b_x, \quad b_y, \quad b_z := \sqrt[+]{1 - b_x^2 - b_y^2} \right]^T \text{ and the point on line}$$

$$\boldsymbol{a} = \alpha_1 \left[1 - \frac{b_x^2}{1+b_z}, -\frac{b_x b_y}{1+b_z}, -b_x \right]^T + \alpha_2 \left[-\frac{b_x b_y}{1+b_z}, 1 - \frac{b_y^2}{1+b_z}, -b_y \right]^T.$$



Unit vectors in the original and rotated coordinate systems, y-axis omitted. The B-plane contains the origin and is normal to the orientation vector $\boldsymbol{b}$. The point-on-line $\boldsymbol{a}$ has planar coordinates (α_1, α_2) in the rotated system.

- Given $n/3$ *observed points* $\boldsymbol{p}_i = [x_i,\ y_i,\ z_i]^T$ with random errors $\boldsymbol{e}_i = [e_{x_i}, e_{y_i}, e_{z_i}]^T$, $i = \{1, 2, \ldots, n/3\}$, for the k-th point, write observation equations:

$$x_k = a_x + t_k b_x + e_{x_k}, \quad y_k = a_y + t_k b_y + e_{y_k}, \quad z_k = a_z + t_k b_z + e_{z_k}$$

- Then a system of n observation equations and two fixed-constraint equations can be represented by a *Gauss-Markov Model (GMM) with constraints:*

$$\underset{n\times 1}{\boldsymbol{y}} = A \underset{m\times 1}{\boldsymbol{\xi}} + \boldsymbol{e}, \qquad \boldsymbol{e} \sim (\boldsymbol{0}, \sigma_0^2 \underset{n\times n}{P^{-1}}), \qquad \operatorname{rk} A = m - 2,$$

$$\boldsymbol{\kappa}_0 = \underset{2\times m}{K} \boldsymbol{\xi}, \qquad \operatorname{rk}\left[A^T \mid K^T\right] = m, \qquad m = 6 + n/3.$$

$$\boldsymbol{\xi} = [a_x, a_y, a_z, b_x, b_y, b_z, t_1, \ldots, t_{n/3}]^T,\ \boldsymbol{\kappa}_0 = \begin{bmatrix} -\boldsymbol{a}^T\boldsymbol{b} \\ 1 - \boldsymbol{b}^T\boldsymbol{b} \end{bmatrix},\ K = [\underset{2\times 6}{K_1} \mid \underset{2\times n/3}{K_2}] = \left[\begin{array}{cc|c} \boldsymbol{b}^T & \boldsymbol{a}^T & 0\ldots \\ \boldsymbol{0}^T & 2\boldsymbol{b}^T & 0\ldots \end{array}\right]$$

- A Lagrangian approach to minimizing the random errors leads to the target function

$$\Phi(\boldsymbol{\xi}, \boldsymbol{\lambda}) := (\boldsymbol{y} - A\xi)^T P(\boldsymbol{y} - A\xi) - 2\boldsymbol{\lambda}^T(\boldsymbol{\kappa}_0 - K\xi) = \underset{\boldsymbol{\xi},\boldsymbol{\lambda}}{\text{stationary}}.$$

- The Euler-Lagrange necessary conditions lead to the normal equations

$$\begin{bmatrix} N & K^T \\ K & 0 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\xi}} \\ \hat{\boldsymbol{\lambda}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{c} \\ \boldsymbol{\kappa}_0 \end{bmatrix}, \quad \text{with} \quad [\underset{m\times m}{N} \mid \boldsymbol{c}] := A^T P[A \mid \boldsymbol{y}], \quad \operatorname{rk} N = m - 2.$$

- The least-squares estimator follows as

$$\hat{\boldsymbol{\xi}} = N_K^{-1}\boldsymbol{c} + N_K^{-1}K^T[KN_K^{-1}K^T]^{-1}[\boldsymbol{\kappa}_0 - KN_K^{-1}\boldsymbol{c}], \text{ with } \underset{m\times m}{N_K} := (N + K^T K), \operatorname{rk} N_k = m.$$

- Using Roberts' (1988) parameterization, the (nonlinear) observation equations for the kth observed point are modified to

$$\begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 - \frac{b_x^2}{1+\sqrt{1-b_x^2-b_y^2}} \\ -\frac{b_x b_y}{1+\sqrt{1-b_x^2-b_y^2}} \\ -b_x \end{bmatrix} + \alpha_2 \begin{bmatrix} -\frac{b_x b_y}{1+\sqrt{1-b_x^2-b_y^2}} \\ 1 - \frac{b_y^2}{1+\sqrt{1-b_x^2-b_y^2}} \\ -b_y \end{bmatrix} + t_k \begin{bmatrix} b_x \\ b_y \\ \sqrt{1-b_x^2-b_y^2} \end{bmatrix} + \begin{bmatrix} e_{x_k} \\ e_{y_k} \\ e_{z_k} \end{bmatrix}.$$

- Then a system of n observation equations (without constraints) can be represented by a *Gauss-Markov Model (GMM):*

$$\underset{n\times 1}{\boldsymbol{y}} = A \underset{m\times 1}{\boldsymbol{\xi}} + \boldsymbol{e}, \quad \boldsymbol{e} \sim (\boldsymbol{0}, \sigma_0^2 \underset{n\times n}{P^{-1}}), \quad \text{rk}\, A = m, \quad m = 4 + n/3,$$

$$\boldsymbol{\xi} = [\alpha_x, \alpha_y, b_x, b_y, t_1, \ldots, t_{n/3}]^T.$$

- The Lagrange target function is given by

$$\Phi(\boldsymbol{e}, \xi, \boldsymbol{\lambda}) := (\boldsymbol{y} - A\xi)^T P(\boldsymbol{y} - A\xi) = \underset{\xi}{\text{stationary}}.$$

- Which leads to the least-squares estimator

$$\hat{\boldsymbol{\xi}} = N^{-1}\boldsymbol{c}, \quad \text{with} \quad [\underset{m\times m}{N} \mid \boldsymbol{c}] := A^T P[A \mid \boldsymbol{y}], \quad \text{rk}\, N = m.$$

- First, assume the conventional six-point-and-orientation parameterization.
- Introduce true variables $\boldsymbol{\mu}_k = [\mu_{x_k}, \mu_{y_k}, \mu_{z_k}]^T$ for the kth measured point $\boldsymbol{p}_k = [x_k, y_k, z_k]^T$, with random errors $\boldsymbol{e}_k$, such that

$$\boxed{\mu_{x_k} := x_k - e_{x_k}, \quad \mu_{y_k} := y_k - e_{y_k}, \quad \mu_{z_k} := z_k - e_{z_k}.}$$

- Then the parametric term t_k can be written in terms of the true variables and the point and slope parameters as

$$\boxed{t_k = \frac{\mu_{x_k} - a_x}{b_x} = \frac{\mu_{y_k} - a_y}{b_y} = \frac{\mu_{z_k} - a_z}{b_z},}$$

which can be used to write two *non-linear condition equations* for the kth data point, thereby eliminating the parametric term t_k and also the columns of zeros in the constraint matrix $K = [K_1 \mid K_2]$

$$\Phi_k(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\mu}_k) = \begin{bmatrix} b_z(\mu_{x_k} - a_x) - b_x(\mu_{z_k} - a_z) \\ b_z(\mu_{y_k} - a_y) - b_y(\mu_{z_k} - a_z) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{\kappa}_0 = \underset{2\times 6}{K_1}\,\boldsymbol{\xi}.$$

- Eliminate two more parameters by adapting Roberts' representation:

$$\Phi_{k,1}(\alpha_1, \alpha_2, b_x, b_y, \boldsymbol{\mu}_k) := \sqrt{1-b_x^2-b_y^2} \cdot \Big\{\mu_{x_k} -$$

$$-\alpha_1\left[1-b_x^2/\left(1+\sqrt{1-b_x^2-b_y^2}\right)\right]-\alpha_2\left[-b_x b_y/\left(1+\sqrt{1-b_x^2-b_y^2}\right)\right]\Big\}-b_x\left[\mu_{z_k}+\alpha_1 b_x+\alpha_2 b_y\right] = 0,$$

$$\Phi_{k,2}(\alpha_1, \alpha_2, b_x, b_y, \boldsymbol{\mu}_k) := \sqrt{1 - b_x^2 - b_y^2} \cdot \Big\{\mu_{y_k} -$$

$$-\alpha_1\left[1+b_x b_y/\left(1+\sqrt{1-b_x^2-b_y^2}\right)\right]-\alpha_2\left[1-b_y^2/\left(1+\sqrt{1-b_x^2-b_y^2}\right)\right]\Big\}-b_y\left[\mu_{z_k}+\alpha_1 b_x+\alpha_2 b_y\right] = 0.$$

- $\Phi_k(\alpha_1, \alpha_2, b_x, b_y, \boldsymbol{\mu}_k) = [\Phi_{k,1}, \Phi_{k,2}]^T$ is nonlinear. Linearize using the Taylor series expansion

$$\Phi_k(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\mu}_k) \approx \Phi_k^0 + d\Phi_{\boldsymbol{\alpha},k}^0 \cdot d\boldsymbol{\alpha} + d\Phi_{\boldsymbol{\beta},k}^0 \cdot d\boldsymbol{\beta} + d\Phi_{\boldsymbol{\mu},k}^0 \cdot d\boldsymbol{\mu}_k + \dots,$$

  where $\boldsymbol{\alpha} := [\alpha_1, \alpha_2]^T$, $\boldsymbol{\beta} := [b_x, b_y]^T$, and the incremental vectors are defined by

$$\boldsymbol{\xi} = [d\boldsymbol{\alpha}^T, d\boldsymbol{\beta}^T]^T := [d\alpha_1, d\alpha_2, db_x, db_y]^T,$$

$$d\boldsymbol{\mu}_k := [x_k - \mu_{x_k}^0 - e_{x_k}, y_k - \mu_{y_k}^0 - e_{y_k}, z_k - \mu_{z_k}^0 - e_{z_k}]^T.$$

- Collect the first-order derivatives in matrices A_k and B_k such that

$$\underset{2\times 4}{A_k} := [-d\Phi_{\boldsymbol{\alpha},k}^0 | -d\Phi_{\boldsymbol{\beta},k}^0] = -\begin{bmatrix} \frac{\partial \phi_{k,1}}{\partial \alpha_1} & \frac{\partial \phi_{k,1}}{\partial \alpha_2} & \frac{\partial \phi_{k,1}}{\partial b_x} & \frac{\partial \phi_{k,1}}{\partial b_y} \\ \frac{\partial \phi_{k,2}}{\partial \alpha_1} & \frac{\partial \phi_{k,2}}{\partial \alpha_2} & \frac{\partial \phi_{k,2}}{\partial b_x} & \frac{\partial \phi_{k,2}}{\partial b_y} \end{bmatrix}, \underset{2\times 3}{B_k} := d\Phi_{\boldsymbol{\mu},k}^0 = \begin{bmatrix} \frac{\partial \phi_{k,1}}{\partial \mu_{x_k}} & \frac{\partial \phi_{k,1}}{\partial \mu_{y_k}} & \frac{\partial \phi_{k,1}}{\partial \mu_{z_k}} \\ \frac{\partial \phi_{k,2}}{\partial \mu_{x_k}} & \frac{\partial \phi_{k,2}}{\partial \mu_{y_k}} & \frac{\partial \phi_{k,2}}{\partial \mu_{z_k}} \end{bmatrix},$$

  together with the vector $\boldsymbol{w}_k := \boldsymbol{\Phi}_k^0 + B_k(\boldsymbol{p} - \boldsymbol{\mu}_k^0)$, which for the kth point leads to the system

$$\boldsymbol{w}_k = A_k \boldsymbol{\xi} + B_k \boldsymbol{e}_k$$

  when higher order terms of the Taylor series are neglected.
- Then, the system for all $n/3$ data points, together with the distribution of the random error vector $\boldsymbol{e}$, can be written as a Gauss-Helmert Model:

$$\boxed{\underset{(2n/3)\times 1}{\boldsymbol{w}} = A \underset{4\times 1}{\boldsymbol{\xi}} + B \underset{n\times 1}{\boldsymbol{e}}, \quad \boldsymbol{e} \sim (\mathbf{0}, \sigma_0^2 \underset{n\times n}{P^{-1}}),}$$

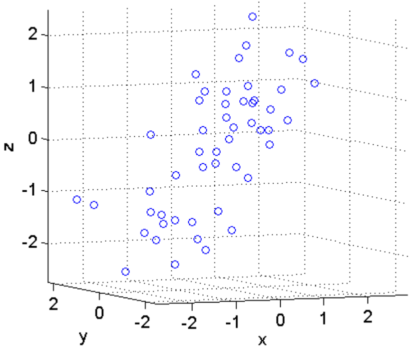
  with model redundancy $r = 2n/3 - 4$ in this example.

- The following Lagrange target function, with $(2n/3) \times 1$ vector of Lagrange multipliers, is defined in order to minimize the random error vector $\boldsymbol{e}$, while satisfying the given model:
$$\Phi(\boldsymbol{e}, \boldsymbol{\xi}, \boldsymbol{\lambda}) = \boldsymbol{e}^T P \boldsymbol{e} + 2\boldsymbol{\lambda}^T(\boldsymbol{w} - A\boldsymbol{\xi} - B\boldsymbol{e}) = \underset{\boldsymbol{e},\boldsymbol{\xi},\boldsymbol{\lambda}}{\text{stationary}}.$$
- The Euler-Lagrange necessary (first-order) conditions are
$$\frac{1}{2}\frac{\partial \Phi}{\partial \boldsymbol{e}} = P\tilde{\boldsymbol{e}} - B^T\hat{\boldsymbol{\lambda}} \doteq 0, \quad \frac{1}{2}\frac{\partial \Phi}{\partial \boldsymbol{\xi}} = -A^T\hat{\boldsymbol{\lambda}} \doteq 0, \quad \frac{1}{2}\frac{\partial \Phi}{\partial \boldsymbol{\lambda}} = \boldsymbol{w} - A\hat{\boldsymbol{\xi}} - B\tilde{\boldsymbol{e}} \doteq 0.$$
- Solving the above system yields the parameter estimator
$$\boxed{\hat{\boldsymbol{\xi}} = \underbrace{\left[A^T(BP^{-1}B^T)^{-1}A\right]}_{4\times 4}^{-1} A' \underbrace{(BP^{-1}B^T)}_{(2n/3)\times(2n/3)}^{-1} \boldsymbol{w}}$$
and the residual predictor
$$\tilde{\boldsymbol{e}} = P^{-1}B^T \cdot \hat{\boldsymbol{\lambda}} = P^{-1}B^T(BP^{-1}B^T)^{-1}(\boldsymbol{w} - A\hat{\boldsymbol{\xi}}).$$
- Applying the law of variance propagation yields the estimated dispersion of $\hat{\boldsymbol{\xi}}$ as
$$\hat{D}\{\hat{\boldsymbol{\xi}}\} = \hat{\sigma}_0^2 \left[A^T(BP^{-1}B^T)^{-1}A\right]^{-1};$$
where $\hat{\sigma}_0^2$ is an estimate for the variance component such that
$$r \cdot \hat{\sigma}_0^2 = \tilde{\boldsymbol{e}}^T P \tilde{\boldsymbol{e}} = \boldsymbol{w}^T(BP^{-1}B^T)^{-1}(\boldsymbol{w} - A\hat{\boldsymbol{\xi}}) = \boldsymbol{w}^T\hat{\boldsymbol{\lambda}},$$
with $r = 2n/3 - 4$ being the system redundancy.
- Both the estimator $\hat{\boldsymbol{\xi}}$ and the predictor $\tilde{\boldsymbol{e}}$ are *unbiased* since $E\{\boldsymbol{w}\} = A\boldsymbol{\xi}$.

- The "Flight of the Bumblebee" data set as shown in Petras and Podlubny (2007).
  - The data come from a random sample of $n/3 = 50$ points drawn from a multivariate normal distribution with

$$\text{mean } \boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ and covariance } \Sigma = \begin{bmatrix} 1 & 1/5 & 7/10 \\ 1/5 & 1 & 0 \\ 7/10 & 0 & 1 \end{bmatrix},$$

generated using MATLAB functions `rng(5,'twister')` and `mvnrnd(`$\boldsymbol{\mu}, \Sigma$`,50)`.

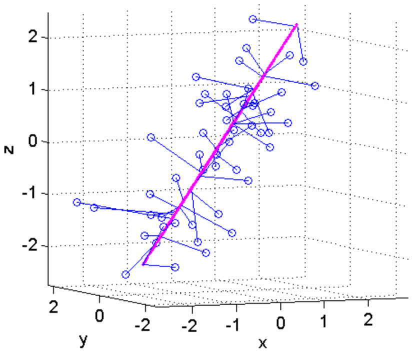


Estimated point $\hat{\boldsymbol{a}}$, orientation vector $\hat{\boldsymbol{b}}$, and their standard deviations derived from $\hat{D}\{\hat{\xi}\}$.

$$\left.\begin{aligned} \hat{\beta} &= \begin{bmatrix} \hat{b}_x & \hat{b}_y \end{bmatrix}^T \\ \hat{\boldsymbol{a}} &= R(\hat{\beta}) \cdot \begin{bmatrix} \hat{\alpha}_1 & \hat{\alpha}_2 & 0 \end{bmatrix}^T \\ \hat{b}_z &:= \sqrt[+]{1 - \hat{b}_x^2 - \hat{b}_y^2} \end{aligned}\right\} \Rightarrow$$

$\hat{a}_x$	-0.067111	± 0.091456
$\hat{a}_y$	0.047785	± 0.121017
$\hat{a}_z$	0.049820	± 0.088503
$\hat{b}_x$	0.677404	± 0.058450
$\hat{b}_y$	0.219309	± 0.077523
$\hat{b}_z$	0.702159	± 0.056575

Estimated variance component: $\hat{\sigma}_0^2 = 0.764288$. Redundancy: $r = 2n/3 - 4 = 96$.



- We have solved the 3D line-fitting problem by Total Least-Squares Adjustment using a minimum parameterization within the Gauss-Helmert Model.
- This approach also allows estimation of a variance component and computation of standard deviations for the parameter estimates, thus permitting statistical hypothesis testing.
- The size of the matrix $[BP^{-1}B^T]^{-1}$ depends on the number of measured points.
  - However, the matrix is 2×2 block-diagonal for uncorrelated observations (i.e., for a diagonal weight matrix P).
  - This also turns out to be the case when there are correlations between individual point coordinates (i.e., for a 3×3 block-diagonal weight matrix P), but not among the points.
- Full details of this work will be presented in a future paper.

I. Petras and I. Podlubny. State space description of national economies: The V4 countries. *Computational Statistics & Data Analysis*, 52(2):1223–1233, Oct. 2007.

A. J. Pope. Some pitfalls to be avoided in the iterative adjustment of nonlinear problems. In *Proceedings of the 38th Annual Meeting of the ASPRS*, pages 449–477, Falls Church, VA, 1972. American Society of Photogrammetry.

K. Roberts. A new representation for a line. In *Proceedings of The Computer Society Conference on Computer Vision and Pattern Recognition*, pages 635–640, Ann Arbor, Michigan, 1988. Computer Society Press.

B. Schaffrin and K. Snow. Total Least-Squares regularization of Tykhonov type and an ancient racetrack in Corinth. *Linear Algebra and its Applications*, 432(8): 2061–2076, Apr. 2010.