

## 2-Dimensional Geometric Analysis of a Simple Free Network

Sluis, B. and C. Toth

Dept. of Civil, Environmental and Geodetic Engineering, The Ohio State University, 470 Hitchcock Hall, 2070 Neil Avenue  
Columbus, OH 43210, USA - (sluis.2, toth.2)@osu.edu

Commission IV WG I/6

**KEY WORDS:** Free-Network, Geometry, Local Reference System, Gauss-Markov/Gauss Mixture Model

### ABSTRACT:

This paper attempts to quantify geometric considerations in observations and observe trends in solutions to free network solutions. The method of investigation will be utilizing 2D observations to determine how each measurement affects the overall solution and the location of the observations relative to the other nodes. A local reference system will be determined using the Gauss-Markov model with constraints by fixing the largest range observation to the y-axis to give a relative orientation. Further solutions will be calculated by fixing additional points to generate multiple least squares solutions relative to the local reference system. The resulting final points will be modeled using the Gauss-mixture model and compared to a simulated dataset generated by adding random error to the observations. Different weight matrices will be tested to demonstrate the effect on the overall solution. These methods were chosen because of prior experimentation by different research groups studying geometric considerations for UAS and ground surveying conditions. The major contribution will be the trends observed in the modeling and the correlation of the fixed local solutions to the geometry of the points.

### 1. INTRODUCTION

#### 1.1 Background

Current surveying techniques emphasize accuracy and cost as the driving concerns for project management. For example, a using photogrammetry UAS system costs more than a lidar scanning per square kilometer to create a DEM. Both systems are heavily studied, and the flight planning, including sensor selection, using ground control and processing methods dictate expected accuracy of the point cloud. However, the effect of the geometric aspects of the surveys in terms of ground and air control are not fully understood. Solutions often utilize excess ground control to avoid geometric considerations, which are vital for accurate detection of centimeter-scale landform change and to reduce the cost of operation (Harwin et al., 2015). Generating cost effective approximations for the error expected in free network solutions will potentially reduce the overall cost of the survey.

Geospatial data solutions produced from UAS surveys involve multiple coordinate systems. Direct geo-referencing is used for initial sensor orientation in a global reference coordinate system. In GPS-denied environment, however, only an arbitrary local system can be used. In general, in situations where the infrastructure is inadequate, a local coordinate system must be used (Dewitt et al., 2000). Indirect geo-referenced systems use measurements of control points in some of the images and fix the local coordinate system to global/mapping system. The ground control used in the processing impacts the accuracy of the overall solution in the final coordinate system used.

Previous studies indicate evenly distributed control points generate the highest accuracy in the final solution but there are many conditions to define 'well-distributed'. The best control configurations utilize sufficient control at each end of the surveyed block and then evenly distributed points inside the area, all of them must be highly visible in as many images as possible, and, if possible, utilize variable elevations (Shahbazi et al., 2015). It is common in these studies to have an excess of ground control and then test solutions with less ground control points. Many times, the points are split into two groups, one used in the computations and the other ones, called check points are used for quality checks. The results of this testing indicate that a large amount of ground control points can be removed without significant losses in accuracy (Mancini et al., 2013).

#### 2. MODEL CONDITIONS

The purpose of this model is to generate a 2D solution that is the most accurate location of the points. The Gauss-Markov model with constraints is the ideal model to represent the nodes because the fixed constraints can be applied to the x and y locations for a specific node. These fixed nodes will be the control/anchor nodes. Remaining nodes will be free nodes and their location will vary based on a least squares estimation. The combination of these nodes produces a local solution.

##### 2.1 Gauss-Markov Model with Constraints

The Gauss-Markov model is used as an observational model for adjustments. The original Gauss-Markov model uses an observational model represented by:

$$y = A\xi + e, rkA =: q \leq \{m, n\} \quad (1)$$

$$e \sim (0, \sigma_0^2 P^{-1}) \quad (2)$$

$$\kappa_0 = K \xi, rk K =: l \geq m - q, rk [AT, KT] = m \quad (3)$$

The left side of the equation  $y$  represents the observations. These are the distance measurements between nodes. The  $A$  matrix is a partitioned Jacobian matrix with derivatives with respect to  $m$ . This matrix is further divided into  $\kappa_0$  which represents the constrained node locations as known values. Since this is a 2D model, the derivatives will be with respect to the x and y location.  $\xi$  represents the unknown correction parameters and  $e$  represents the random error. The distribution of the errors is assumed to be Gaussian based on the weight matrix,  $P$  (Schaffrin and Snow, 2017). Assuming the observations will always be larger than the rank of the Jacobian matrix means that there will be redundant observations. This means that there is a least squared solution.

The data utilized in this study are observations of distances between points. The initial setup included ranges between each vertex and observations in both directions (i.e., 10 to 3 and 3 to 10, see Section 2.2). These ranges define a free network without a defined global solution. So, computing a local solution is the best that can be done. The free network can be arbitrarily oriented, and the direction of the largest variation is oriented along the y coordinate axis. These geometric variances include

scale of observations from 700 to 30 meters, uneven point distributions relative to centroid of the network, and 5 of the 10 points being collinear. The varying ranges will be tested in the weight matrix. Collinear points are used when comparing the RMSE.

## 2.2 Gaussian Mixture Model (GMM)

Gauss-Markov solutions yield 10 individual points generated in a local coordinate system as seen in Fig. 1. The lines connecting each point demonstrate 2 unique observations: two independent range measurements from both points. The RMSE was minimized for the local reference solution by fixing the largest observed distance, which observation was between points 9 and 2. Fig. 1 represents the reference local solution. A good geometry for control in this solution will utilize the furthest points that are not collinear with other points. This point combination would include a pairing of 2, 9, and 4 or something similar. This combination gives the largest area inside the figure.

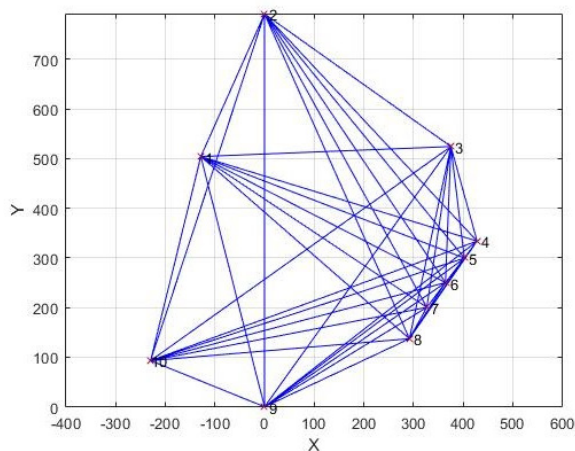


Figure 1: Network setup with connected observations (meters)

Multiple localized reference solutions were calculated fixing different points. Each of these solutions fixed 3 of the 10 nodes from the solved 2-9 solution. These multiple local solutions are the 3-constraint solutions. Each of these 3-constraint solutions minimizes the RMSE when compared to the 2-9 solution. The solutions used the same observations and gave the location of each free node. An example of the distribution of the 3-constraint local solution can be seen in Fig. 2. These multiple solutions demonstrate the lowest RMSE local 3-constraint solutions around point 2. The black point represents the 2-9 solution while each blue cross represents a unique 3-constraint solution.

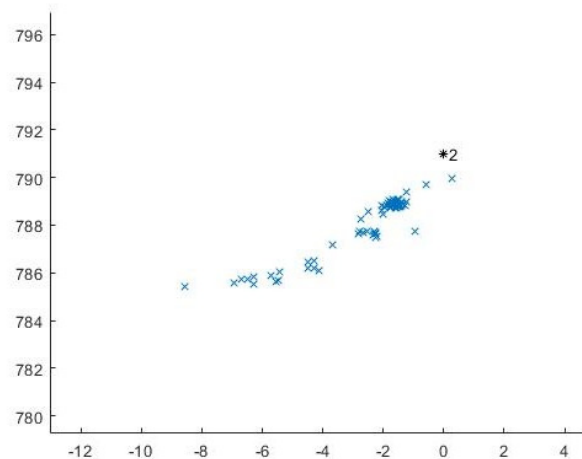


Figure 2: Point 2 lowest RMSE local solutions

Using the local 3-constraint solutions from each of the 120 combinations, the GMM takes the resulting vertices from the multiple solutions and clusters the data around a central point. This is a method for determining the distribution of the resulting data. Each point is assumed to have a local reference system value that is unknown. The GMM algorithm determines the most likely location of that local reference system based on the distribution of the estimates from the multiple fixed points (He et al., 2011). However, this GMM needs defined properties for the weight matrix based on the distribution.

Within the GMM, there is a single covariance matrix with different possible properties. The first is a diagonal covariance matrix that says the predicted values are uncorrelated. Obviously, this is not the case because certain mutual solutions have common control node. Ultimately this yields a circular or oval shape symmetrical in the x and y axis. A full covariance matrix allows for each point to be related to each other and thus providing the best fit while not tying the values to symmetry from the x and y axis. This result, however, can lead to overfitting the dataset because the ellipse can vary to incorporate all points in a single circle with Gaussian distribution.

Covariance matrices are either shared or unshared between solutions. A shared covariance matrix implies that all the covariance matrices are the same and the distribution of the ellipses is indicative of the same orientation. Unshared covariance matrices for each solution change the orientation of the Gaussian ellipse. This means the orientation of each ellipse is not perpendicular to the x and y axis (Section 3.2).

The nature of the problem indicates that the free nodes will be correlated. This is because each solution uses 3 fixed points, and these points are shared between solutions. A full unshared covariance matrix is the best indication for the solution because the correlation between solutions will be related to the geometry of the system (He et al. 2011). The geometry of the system solutions will be independent of the assigned axis because the orientation of the points will dictate the variation in the solution.

### 3. SIMULATION RESULTS

#### 3.1 Cluster Modeling Results

The next simulation was to test the overall location of the free nodes and determine the effect of the spatial distribution on the overall solution quality. The free nodes in the 3-constraint local solutions did not appear to maintain a uniform Gaussian distribution around the 2-9 constraint solution. The points were not in a circular cluster around the 2-9 constraint solution. Using the model defined, the residual matrix for each 3-constraint solution are solved when compared to the 2-9 solution. Standard practice is to assume the variation in the random error has a normal distribution and is centered at 0 (Schaffrin and Snow, 2017). Multiple weight matrices are tested against the distribution. The original local reference system solution utilized a diagonal weight matrix giving equal weight to all observations. This is standard for most applications (Schaffrin and Snow, 2017). A length-based diagonal weight matrix was used for the next round of simulations. A length-based matrix utilizes the range between nodes to account for geometric considerations and error increasing with distance. A larger distance in an observation normally indicates a larger error. Another weight matrix considered utilized the shared points. Solutions which share any amount of points have a sparse covariance matrix. If a single point was shared between the solutions, a fixed value of 0.33 was allocated to the shared point. An example is the 3-constraint solution using points 2-9 and 4 when compared with a 3-constraint solution using 2-6 and 3 as fixed nodes was given a value of 0.33. If two points were shared, a fixed value of 0.67 was allocated. The locations of the non-diagonal values in the matrix align with the shared fixed points between the solutions. The purpose of this modeling is to generate an accurate weight matrix to represent the data distribution.

#### 3.2 Gaussian Mixture Modeling Results

The Figs. 3 and 4 below represent the different covariance matrices tested on point 2. The top two graphs represent the results using a diagonal covariance matrix to fit the dataset. These matrices assume there is a uniform Gaussian distribution with a centroid at the original 2-9 constraint solution. Fig. 3 demonstrates Gaussian circles centered on the local reference system that have a diagonal covariance matrix do not accurately support the data distribution. The bottom two images demonstrate a sparse covariance matrix is more representative of the distribution of solutions. Since a full sigma matrix more accurately represents the dataset, the covariance matrix shows the points are geometrically related. The next experiment involves clustering the data using supervised and unsupervised fitting.

The top two graphs utilized individual fitting and the algorithm determined two circles best fit the dataset. This data fitting for 87 observations fit a normal distribution with 16 solutions outside the dataset. The bottom graph represents the optimal data fitting using a full and shared covariance matrix. Here the Gaussian circles were fitted in a similar manner to the shared covariance matrix but were unsupervised, so the number of groupings was undefined. The unsupervised orientation utilized two groups for the dataset which indicates the solutions can be partitioned into multiple groups to optimally represent the spread of the dataset. This means the data cannot fit a single Gaussian distribution and since the data originates from a single dataset, the geometry of the solutions affects the Gaussian distribution.

Testing indicated that the geometric location of the 3 constrained nodes affected the quality of the solutions. These solutions were modeled in the Gaussian mixture model using a diagonal and sparse covariance matrix. The diagonal covariance matrix with a circular Gaussian distribution did not properly model the data. A sparse covariance matrix utilizing the similar points between solutions fit the data. The unsupervised classification of the Gaussian mixture model indicated the local solutions did not fit a single Gaussian distribution. This further indicated a geometric relationship is present in the 3-constraint local solutions. The next set of testing is to compare the original 3-constraint local solutions with a noisy dataset and resulting 3-constraint solutions. A dataset with added noise that has similar trends in the solutions will demonstrate the properties of optimal geometric local solutions and remove the bias of a single dataset.

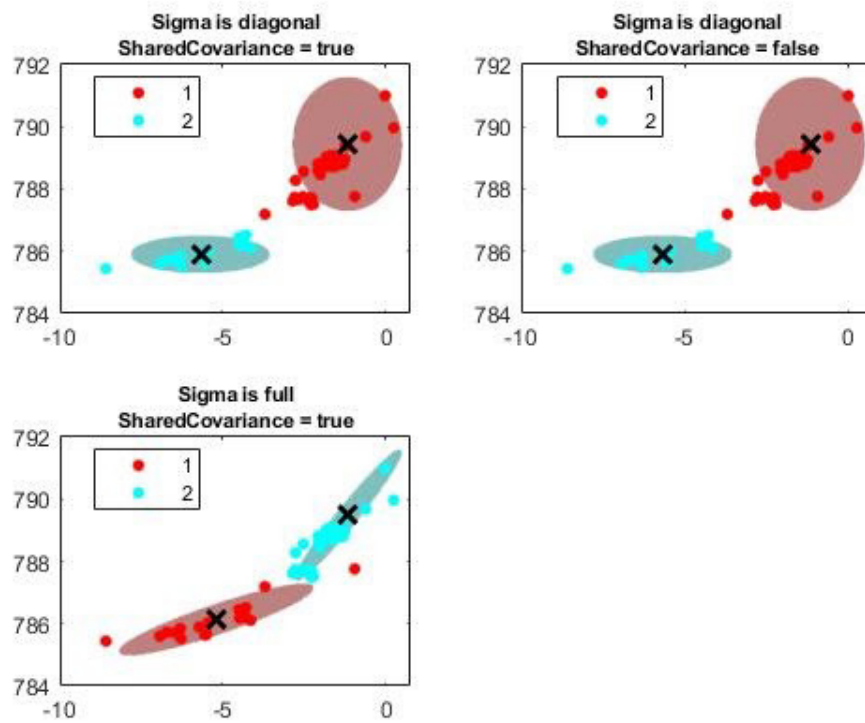


Figure 3: Sigma distribution with diagonal and full covariance matrices

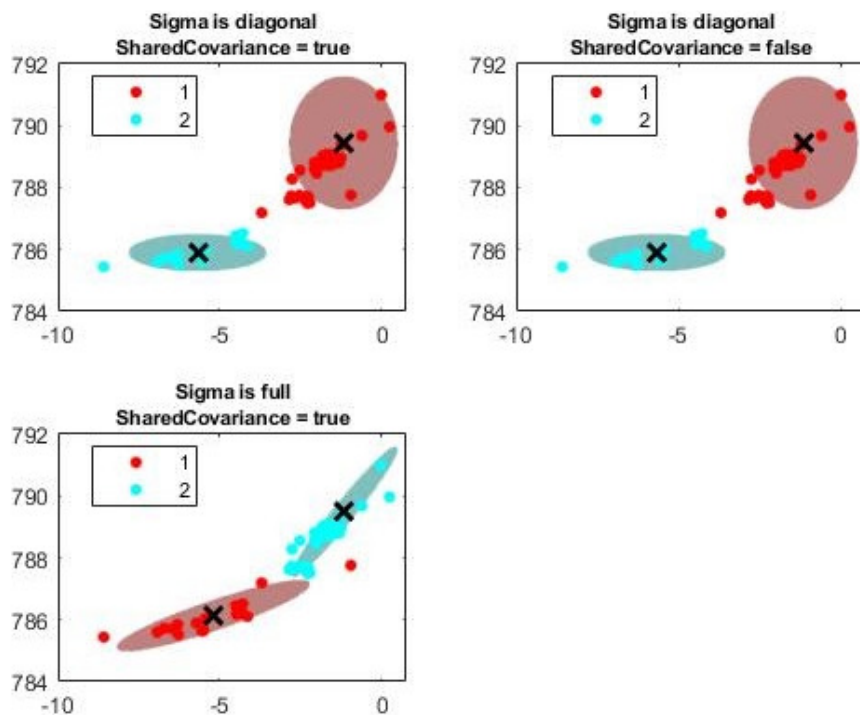


Figure 4: Gaussian mixture model with unsupervised sorting

#### 4. NOISY DATASET AND RMSE RESULTS

##### 4.1 Simulation Dataset Results

If a new dataset is generated that exhibits the same trends as the original dataset, then it can be assumed the observations and geometric considerations can be applied to other networks. Each range observation was given a variance of 5 meters to generate a new dataset to compare against the 2-9 constraint solution. This data was processed as the noisy dataset. The RMSE data from the original 3-point local solutions and the new noisy dataset can be seen below. The fixed points are organized from least to greatest then plotted on the graph with the x axis ranging from 1 to 8, y axis from 2 to 9, and the z axis from 3 to 10. This represents all the unique combinations of the 3 control points from a subspace of 10 points. Any line parallel with an axis represents a solution that has 2 points in common. An example is the bubbles on the right side of each graph all share the points 1 and 2 with a varying third point. This applies to columns as well with the farthest left column representing solutions with 9 and 10 as shared points. The larger bubbles represent a higher RMSE value. Values in the title of the graph represent the multiplication of the RMSE for all values to generate the bubble graph.

The data demonstrates correlations between any two points. The original dataset and the noisy data have nearly identical RMSE when properly scaled. This proves that the RMSE correlation is invariant of the noise in the dataset. This leaves only the geometric considerations to compare. The highest values from the graph are from the solution fixing points 7, 8 and 9. The other peak values come from solutions involving any combinations of 4 through 9. The points 7, 8, and 9 are close to collinear. This local solution, when connected, also creates a figure that does not encompass any of the other points. This would be defined as a poor geometry. The lowest value for RMSE was the solution with 2, 3 and 9. Selecting these three points provides the furthest two points and a third point that is not collinear with the non-selected points while also being closest to most of the points. This is the expected outcome and would indicate a traditional good geometry for the figure including points being spread and not collinear.

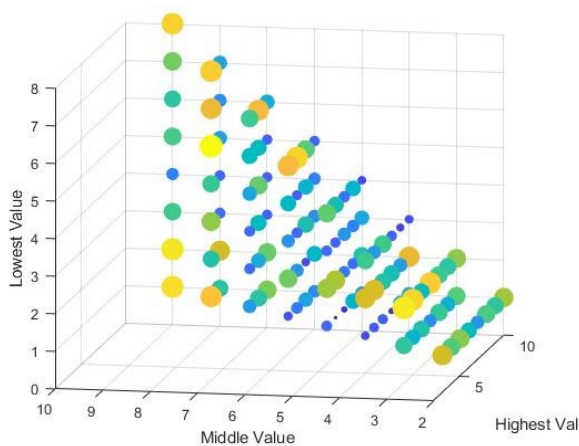


Figure 5: RMS value of variance dataset (50x)

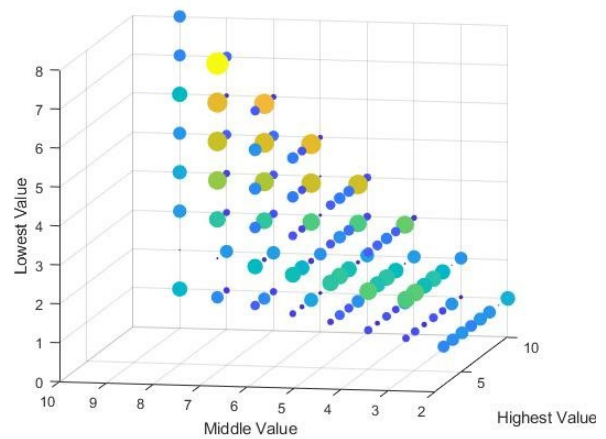


Figure 6: RMS value of base dataset (100x)

#### 5. CONCLUSION

The purpose of this experimentation was to demonstrate a significant correlation between the geometry of generated local solutions and the quality of the local solutions. This was demonstrated using a multifaceted analysis of a free network containing 10 points. A local solution was generated from fixing the furthest 2-points to the y axis. Further local solutions were calculated by fixing 3 points in every combination from the dataset. Then, a noisy dataset was created, and the processes were repeated. The initial solution was compared with the base and noisy data to confirm the trends observed.

The purpose of this paper is to study a simple 10-point 2-dimensional free network to model geometric considerations. The first method considered was the Gaussian Mixture Model. GMM solution utilized multiple different weight matrices. The covariance matrices involved diagonal and full matrices with weights using the range and collinearity. Results of the GMM indicated that a sparse matrix was ideal for modeling because the solutions were related. This is expected because the solutions share fixed nodes, so they are geometrically related. Next, a new dataset was produced from the initial dataset by adding 5 meters of variance in the measurements. A new set of 3-constraint solutions are calculated and compared with the original solution using RMSE. The original data local solutions and the new noisy local solutions have the same correlation between points demonstrating the relationship between geometry and RMSE. The results of the Gaussian mixture model and the RMSE analysis demonstrate geometric considerations have large effects on the fixed solution quality. After studying the simple free network model, the considerations for an optimal geometry used in selecting control should be proximity to other points, the largest distance between points, and selecting points that are not collinear with more than 1 other point.

Future work for this project will quantify the geometric considerations. The trends can be quantified with multiple initial solutions using variable geometry. This could include more points of using a convex geometry with internal points. Another consideration will be to utilize 3 dimensional solutions to create a more accurate representation of a field experiment.

## REFERENCES

- DeWitt, B.A.; Wolf, P.R. *Elements of Photogrammetry (with Applications in GIS)*, 3rd ed.; McGraw-Hill Higher Education: New York, NY, USA, 2000
- Harwin, S.; Lucieer, A.; Osborn, J. 2015. *The impact of the calibration method on the accuracy of point clouds derived using unmanned aerial vehicle multi-view stereopsis.*
- He, X., Cai, D., Shao, Y., Bao, H., Han, J.: *Laplacian regularized gaussian mixture model for data clustering.* IEEE Trans. Knowl. Data Eng. 23, 1406–1418 (2011)
- Mancini, F.; Dubbini, M.; Gattelli, M.; Stecchi, F.; Fabbri, S.; Gabbianelli, G. 2013. *Using Unmanned Aerial Vehicles (UAV) for High-Resolution Reconstruction of Topography: The Structure from Motion Approach on Coastal Environments.* Remote Sens.
- Schaffrin and Snow. 2017: *Notes on Adjustment Computations.*
- Shahbazi, M.; Sohn, G.; Théau, J.; Menard, P. 2015. *Development and evaluation of a UAV-photogrammetry system for precise 3D environmental modeling.*