Magical Least Squares:
or
When Is one least squares adjustment better than another?

Earl F. Burkholder, PS, PE
NMSU Department of Surveying Engineering
Las Cruces, NM 88003

Introduction

Least squares, i.e., the sum of the squares of the residuals will be a minimum, has been proven and accepted for adjusting survey data. Least squares is simply the best method available.

Within reason, it is also true that, depending upon how weights are selected, you can get any answer you want using least squares. Therefore, the issue in discussing "How good are my results?" switches from the choice of the tool (least squares) to how the tool was used. Of course, the input data must first be checked and verified blunder-free.

Given blunder-free survey data and a specific statement of how weights are selected, all least squares packages should provide the same answers. Differences from one brand software to another will have to do with the survey data input (formats, weights etc) and what information is included in the report after the adjustment is completed. This article looks at 3 different weighting assumptions on a small network and compares the results.

The example used in this paper is a GPS network based upon two A-order HARN points. Station "Reilly" is located in the central horseshoe of the NMSU campus and Station "Crucesair" is located at the Las Cruces airport some 16 kilometers west of campus. The network consists of 7 independent baselines connecting 4 additional points to the existing HARN stations as shown in Figure 1.

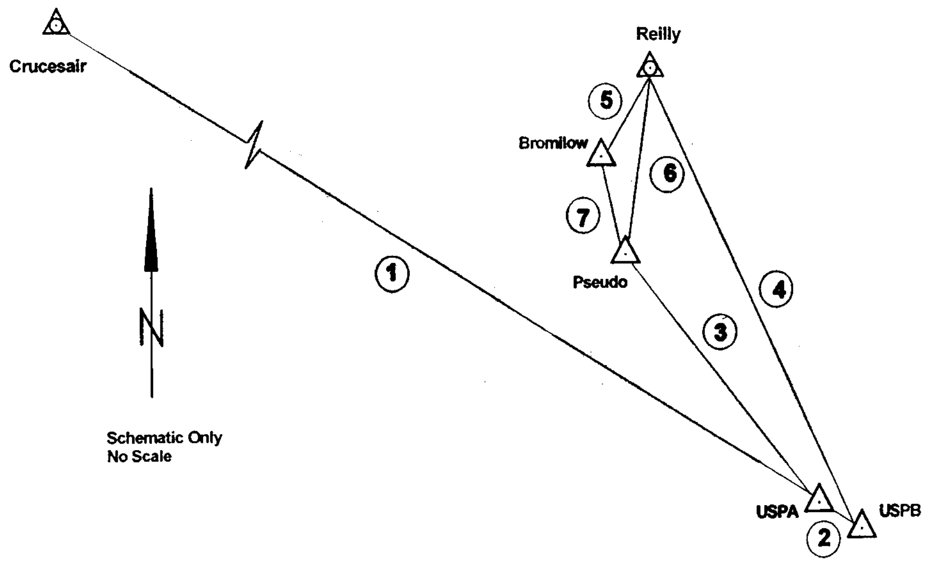

Figure 1 GPS Network at NMSU

The GPS baselines shown and used were collected on three different dates over a period of 4 years. These are not the only baselines on campus nor are they the only observations between the points in question. These baselines were selected because they show excellent consistency, are independent, and include often used points. As explained at the end, there is also a problem (lesson to be learned) with this network.

Control Values and Observed Vectors

The NAD83 geocentric X/Y/Z coordinates for A-order HARN stations "Reilly" and "Crucesair" are as published by the National Geodetic Survey (NGS) and were held fixed in this exercise. They are:

Station Reilly	Station Crucesair
X = -1,556,177.615 m	X = -1,571,430.672 m
Y = -5,169,235.319 m	Y = -5,164,782.312 m
Z = 3,387,551.709 m	Z = 3,387,603.188 m

Single frequency Trimble GPS units were used to collect static data, 57 minutes being the shortest observation time for any of the 7 baselines. The baseline components and the covariance matrix for each observed baseline are as determined by Trimble software using default processing parameters are:

Vector 1 – Crucesair to USPA (use subscript CA):

			Sxx	Syy	Szz
ΔX_{CA} =	15,752.080 m	Sxx	6.321492E-06		
ΔY_{CA} =	-5,179.102 m	Syy	1.545948E-05	4.739877E-05	
ΔZ_{CA} =	-903.089 m	Szz	-1.061303E-05	-3.184780E-05	2.388036E-05

Vector 2 – USPA to USPB (use subscript AB):

			Sxx	Syy	Szz
ΔX_{AB} =	14.971 m	Sxx	1.753823E-07		
ΔY_{AB} =	-15.380 m	Syy	4.411366E-07	1.386877E-06	
ΔZ_{AB} =	-16.660 m	Szz	-3.002835E-07	-9.232852E-07	6.779702E-07

Vector 3 – USPA to Pseudo (use subscript AP):

			Sxx	Syy	Szz
ΔX_{AP} =	-528.036 m	Sxx	9.505016E-08		
ΔY_{AP} =	560.657 m	Syy	8.957064E-08	3.729339E-07	
ΔZ_{AP} =	585.897 m	Szz	-5.022282E-08	-2.221975E-07	3.363763E-07

Vector 4 – USPB to Reilly (use subscript BR):

			Sxx	Syy	Szz
ΔX_{BR} =	-514.003 m	Sxx	3.650165E-07		
ΔY_{BR} =	741.438 m	Syy	9.024127E-07	2.796189E-06	
ΔZ_{BR} =	868.293 m	Szz	-6.189027E-07	-1.881145E-06	1.410196E-06

Vector 5 – Bromilow to Reilly (use subscript MR):

			Sxx	Syy	Szz
ΔX_{MR} =	32.134 m	Sxx	2.762550E-07		
ΔY_{MR} =	51.175 m	Syy	3.200312E-07	6.870545E-07	
ΔZ_{MR} =	94.198 m	Szz	-2.008940E-07	-4.006259E-07	4.661596E-07

Vector 6 – Pseudo to Reilly (use subscript PR):

			Sxx	Syy	Szz
ΔX_{PR} =	29.000 m	Sxx	1.325760E-07		
ΔY_{PR} =	165.422 m	Syy	1.317165E-07	5.265054E-07	
ΔZ_{PR} =	265.719 m	Szz	-7.253348E-08	-3.020965E-07	5.006575E-07

Vector 7 – Bromilow to Pseudo (use subscript MP):

			Sxx	Syy	Szz
ΔX_{MP} =	3.136 m	Sxx	3.367818E-07		
ΔY_{MP} =	-114.242 m	Syy	3.937476E-07	8.766570E-07	
ΔZ_{MP} =	-171.527 m	Szz	-5.186521E-07	-8.977932E-07	1.446501E-06

Blunder Checks

In order to verify the absence of blunders in all baselines, misclosures are computed for each component (X/Y/Z) as follows:

Traverse including baselines 1, 2, and 4 (from "Crucesair" to "Reilly"):

	X	Y	Z
Station Crucesair	-1,571,430.672 m	-5,164,782.312 m	3,387,603.188 m
Baseline 1	15,752.080 m	-5,179.102 m	-903.089 m
Baseline 2	14.971 m	-15.380 m	-16.660 m
Baseline 4	-514.003 m	741.438 m	868.293 m
Computed value	-1,556,177.624 m	-5,169,235.356 m	3,387,551.732 m
Station Reilly	-1,556,177.615 m	-5,169,235.319 m	3,387,551.709 m
Misclosures	**-0.009 m**	**-0.037 m**	**0.023 m**

Loop including baselines 2-3-7-5-4 (being careful to preserve sign convention):

Baseline 2	-14.971 m	15.380 m	16.660 m
Baseline 3	-528.036 m	560.657 m	585.897 m
Baseline 7	-3.136 m	114.242 m	171.527 m
Baseline 5	32.134 m	51.175 m	94.198 m
Baseline 4	514.003 m	-741.438 m	-868.293 m
Misclosures	**-0.006 m**	**0.016 m**	**-0.011 m**

Loop including baselines 5-6-7 (being careful to preserve sign convention):

Baseline 5	32.134 m	51.175 m	94.198 m
Baseline 6	-29.000 m	-165.422 m	-265.719 m
Baseline 7	-3.136 m	114.242 m	171.527 m
Misclosures	**-0.002 m**	**-0.005 m**	**0.006 m**

All baselines have been included in the checks and all misclosures are acceptable. Therefore, it is legitimate to perform a least squares adjustment of the 7 baselines to determine the "best" adjusted position for points USPA, USPB, Pseudo, and Bromilow. Any adjustment should also provide information on the quality of the answers, i.e., "What is the standard deviation of the computed position?" - in both the geocentric (X/Y/Z) reference frame and in the local (east/north/up) reference frame. This paper uses 3 different weighting schemes and shows a comparison of the various answers.

Comparison of Weighting Options on Results of Least Squares Adjustment of GPS Network

Earl F. Burkholder, PS, PE - August 2005

The comparison shown reflects different weighting decisions on the results of the NMSU GPS network based upon A-order HARN stations "Crucesair" and "Reilly." Four new points were surveyed and included as part of the least squares adjustment of the 7 baselines.

Three weighting options are included - equal weights, weights by standard deviations of the baseline components, and weights based upon the full covariance matrix of each baseline.

Geocentric Coordinates and sigma | **Geodetic Coordinates and local sigma**

Option 1, USPA - equal weights

				deg	min	sec		
X =	-1,555,678.5860 m	+/- 0.0063 m	φ =	32	16	23.00014	+/- 0.0063 m	(N)
Y =	-5,169,961.4000 m	+/- 0.0063 m	λ =	106	44	48.90838	+/- 0.0063 m	(E)
Z =	3,386,700.0910 m	+/- 0.0063 m	h =		1178.021 m		+/- 0.0063 m	(U)

Option 2, USPA - weights based upon standard deviation of each baseline component.

				deg	min	sec		
X =	-1,555,678.5800 m	+/- 0.0020 m	φ =	32	16	23.00019	+/- 0.0040 m	(N)
Y =	-5,169,961.3940 m	+/- 0.0043 m	λ =	106	44	48.90822	+/- 0.0023 m	(E)
Z =	3,386,700.0880 m	+/- 0.0040 m	h =		1178.013 m		+/- 0.0041 m	(U)

Option 3, USPA - weights based upon the full covariance matrix of each baseline.

				deg	min	sec		
X =	-1,555,678.5840 m	+/- 0.0026 m	φ =	32	16	23.00016	+/- 0.0052 m	(N)
Y =	-5,169,961.3930 m	+/- 0.0061 m	λ =	106	44	48.90839	+/- 0.0030 m	(E)
Z =	3,386,700.0870 m	+/- 0.0049 m	h =		1178.013 m		+/- 0.0056 m	(U)

Option 1, USPB - equal weights

				deg	min	sec		
X =	-1,555,663.6130 m	+/- 0.0071 m	φ =	32	16	22.36246	+/- 0.0071 m	(N)
Y =	-5,169,976.7680 m	+/- 0.0071 m	λ =	106	44	48.19143	+/- 0.0071 m	(E)
Z =	3,386,683.4240 m	+/- 0.0071 m	h =		1177.917 m		+/- 0.0071 m	(U)

Option 2, USPB - weights based upon standard deviation of each baseline component.

				deg	min	sec		
X =	-1,555,663.6100 m	+/- 0.0023 m	φ =	32	16	22.36247	+/- 0.0062 m	(N)
Y =	-5,169,976.7680 m	+/- 0.0071 m	λ =	106	44	48.19132	+/- 0.0030 m	(E)
Z =	3,386,683.4240 m	+/- 0.0059 m	h =		1177.916 m		+/- 0.0066 m	(U)

Option 3, USPB - weights based upon the full covariance matrix of each baseline.

				deg	min	sec		
X =	-1,555,663.6120 m	+/- 0.0032 m	φ =	32	16	22.36244	+/- 0.0069 m	(N)
Y =	-5,169,976.7680 m	+/- 0.0086 m	λ =	106	44	48.19132	+/- 0.0039 m	(E)
Z =	3,386,683.4230 m	+/- 0.0063 m	h =		1177.916 m		+/- 0.0078 m	(U)

But, these values are defective! The physical distance mark-to-mark = 27.162 m

Comparison of Weighting Options on Results of Least Squares Adjustment of GPS Network

Earl F. Burkholder, PS, PE - August 2005

But, what is the local direction and distance between USPA and USPB based upon Options 2 & 3?

Option 2 - Inverse USPA to USPB

USPA	X =	-1,555,678.5800	USPB	X =	-1,555,663.6100	ΔX =	14.9700	m
	Y =	-5,169,961.3940		Y =	-5,169,976.7680	ΔY =	-15.3740	m
	Z =	3,386,700.0880		Z =	3,386,683.4240	ΔZ =	-16.6640	m

								Network Accuracy	Local Accuracy
Δe =	18.7650	m	Dist =		27.1687 m			+/- 0.0058 m	+/- 0.0040 m
Δn =	-19.6473	m							
Δu =	-0.0974	m	Azi =	136	18	56.58		+/- 43.2 sec	+/- 29.93 sec

Option 3 - Inverse USPA to USPB

USPA	X =	-1,555,678.5840	USPB	X =	-1,555,663.6120	ΔX =	14.9720	m
	Y =	-5,169,961.3930		Y =	-5,169,976.7680	ΔY =	-15.3750	m
	Z =	3,386,700.0870		Z =	3,386,683.4230	ΔZ =	-16.6640	m

								Network Accuracy	Local Accuracy
Δe =	18.7672	m	Dist =		27.1704 m			+/- 0.0028 m	+/- 0.0013 m
Δn =	-19.6475	m							
Δu =	-0.0971	m	Azi =	136	18	45.55		+/- 20.79 sec	+/- 9.65 sec

Notes:

1. The concept of "local accuracy" and "network accuracy" is described in a paper, "Spatial Data Accuracy as Defined by the GSDM," and published in the ACSM Journal of Surveying & Land Information Systems, Vol. 59, No. 1, March 1999.

2. Intuitively, the difference between "local" accuracy and "network" accuracy addresses the question, "Accurcy with respect to what?"

3. Standard deviations of the components are closely related to network accuracy while local accuracy is influenced more heavily by the off-diagonal elements of the covariance matrx.

4. There is a lot of research to do yet on network and local accuracies, but the GSDM handles those concepts with aplomb.

5. The results of the network adjustment described here are deficient because......?