Reduction of the condition number by ridge-type estimation methods in GPS equations systems

P. Abad*

(*) Department of Cartography and Grafic Expresion
University of Las Palmas de Gran Canaria. Spain

Keywords: GPS, ridge-type, condition number.

Abstract

We present an analysis of ridge-type estimation methods and their performance in response to different kinds, ill-conditioned, of equations systems derived from GPS observations, in order to overcome the effects of this ill-conditioning. We deal with the elements of the main diagonal of the normal matrix in order to alter the solution. We compare the solution accuracy with the least squares standard techniques. This method would decrease the spread on the range of the eigenvalues and improve the conditioning of the matrix. Therefore, the solution of GPS equation systems will be also improved.

1. Introduction

The simultaneous observation of four or more satellites of known position from an unknown place enables us to build an equation system using the distance formula. There are two kinds of equations depending on the satellite signal: *phase* equations and *pseudo-range* equations. Throughout this work, we will focus on the pseudo-range equation.

The GPS observation equation following a mechanistic model, [5], [4], after removing the systematic errors can be written in matrix form as follows:

$$Ax + \varepsilon = b \tag{1}$$

Where b is a (m×1) vector of residual observations, A is a (m×n) design matrix, x is a (n×1) vector of corrections to the unknown parameters (dx, dy, dz, dt), and ε is the random error distributed as N (0, σ^2). The partial derivatives of the observation equations with respect to the two types of parameters, station position (dx, dy, dz) and station clock (dt), form the design matrix A.

The least squares solution with a *a-priori* weights R, and *a-priori* parameters (X_0,Y_0, Z_0, t_0) are given by:

$$\hat{x} = (A^t PA)^{-1} A^t Pb \tag{2}$$

The Gauss-Markov theorem says: "*Under the conditions of model (1) the least squares estimator $\hat{x}$ de x is unbiased and have minimum variance among all unbiased linear estimator*"

2. Multicollinearity and influential observations

Least squares results can be sensitive to rounding of data in intermediate stages of calculations. Roundoff errors tend to enter into least squares calculations primarily when the inverse A^tA is taken, (the matrix P can be approximated by the identity); the roundoff errors are particularly great when:

1. A^tA has a determinant, which is close to zero.
2. The element of A^tA differs substantially in order of magnitude.

A solution for the second condition is to transform the variables and thereby reparameterize the regression model. The transformation is called "correlation transformation". It makes all entries in the A'A normal matrix for the transformed variables fall between -1 and +1 (normalized matrix), so that the calculation of the inverse matrix becomes much less subject to roundoff error. This will reduce the condition number of the normal matrix and improve the numerical aspect of the computing procedures. The off-diagonal terms will be less than or equal to one and will represent a quantitative measure of the dependence among model parameters.

The procedure for placing the equations (1) in correlation form begins with the calculation of the matrix D, defined as:

$$D_{ii}^{i} = \left(A^{t} P^{-1} A\right)_{ii}^{-1/2} \tag{3}$$

Where D is a matrix of the diagonal terms of D_{ii} (root square of diagonal covariance matrix element). The relationship between the normalized solution, $\hat{x}_N$ and the standard $\hat{x}$ solution is given as [1]:

$$\hat{x}_N = D^{-1}\hat{x} \tag{4}$$

The equations (1) in correlation form may then be written as:

$$\begin{aligned} b &= ADD^{-1}x + \varepsilon \\ b &= A_N x_N + \varepsilon \end{aligned} \tag{5}$$

This normalized system of equations has the least squares solution:

$$\hat{x}_N = (A_N^t A)^{-1} D A^t P^{-1} b \tag{6}$$

Where

$$A_N^t A = D A^t P^{-1} A D \tag{7}$$

In the first condition the data are nearly multicollinear and this means that the least squares estimates are unstable, the A^tA matrix is near singular and small changes in the observations may result in large changes in the estimates of the parameters. Additionally Silvey [5] and Strang [3] were able to show that the multicollinearity will amplify the effects of observation error on the solution. One measure of the multicollinearity is the condition number of the matrix A^tA or the normalized matrix $A_N^t A_N$. Ridge estimators have been suggested as alternatives to the least squares rule, this method would decrease the spread on the range of the eigenvalues and improve the conditioning of the matrix.

3. Biased estimation with nearly multicollinear data

In the presence of a design matrix A that is nearly collinear, small changes in the values of b, the vector of observations on the dependent variable, may results in dramatic changes in the values of $\hat{x}$, the unbiased least squares estimator of x

Estimators that are not subject to such extreme dependence are a class of biased estimators, known as Ridge-Type estimators; these estimators augment the normalized matrix in a correlation form as:

$$\hat{x}_c = (A_N^t A_N + cI_K)^{-1} D A^t R^{-1} b \tag{8}$$

Where the diagonal matrix cI_K is used as the biasing matrix and c is a constant. In this family of estimators the replacement matrix $A_N^t A_N + cI_K$, which replaces $A_N^t A_N$ in the least squares estimators, has a lower condition number . These estimators have the property that their mean squared error is less tan of $\hat{x}_N$ for a properly chosen $c > 0$, and are more stable than the least squares estimator.

A generalized ridge estimator may be defined as:

$$\hat{x}_K = (A_N^t A_N + K)^{-1} D A^t R^{-1} b \tag{9}$$

Where K is a diagonal positive definite matrix is used as the biasing matrix. The minimum mean square error (MSE) is used for the determination of each diagonal element of K.

$$MSE = E\left[(\hat{x}_K - x_N)^t (\hat{x}_K - x_N)\right] \tag{10}$$

Where $\hat{x}_K$ is the normalized ridge estimate with the bias included and x_N is the true normalized solution.

The optimal choice of K can be obtained by minimizing the MSE function with respect to each k_i, substituting for the solution in a correlation form; one may obtain an equation for the MSE as [2]:

$$MSE = V(\hat{x}_K) + \left[B(\hat{x}_K)\right]^2 \tag{11}$$

Where V and B refer to the total variance and square of the bias respectively.

To obtain a simpler expression for k_i we assume that k_is are equal to k:

$$k = \frac{n}{(\hat{x}_N)^t D^{-2} \hat{x}_N} \tag{12}$$

4. Experimental description

In order to study the reduction of condition number by the ridge-type estimation method we used next data set of satellites observations to determinate the station position and the receiver clock in which the true solution is known. This true solution will allow for accuracy comparisons to be made with the standard least squares method. These observations were provided for the European Space Operation Centre (ESOC) and they are refereed on ALGONQUIN station (Canada) the 2002/9/6 and the number of satellites observed was eight.

The table 1-2, shows the experimental conditions and k:

Data set of observation	Time interval	Processed epoch	Resolution method	Equations × unknown	Number of systems solved
D1	$0^h52^m30^s$	105	LS/RT	8×4	105
D5	$0^h52^m30^s$	105	LS/RT	40×8	21
D10	$0^h\ 89^m$	179	LS/RT	80×13	17
True solution: X = 918129.742 Y = -4346072.212 Z= 4561978.390 t=0.319285e-6					

Table 1 : Experimental description

Data set of observation	Minimum k	Maximum k
D1	0.08307992866562	0.00459407261578
D5	0.01028815556317	0.00282513981236
D10	0.00356952287837	0.00922431102922

Table 2 : Biases

In the resolution we have done the next corrections:

1. Ionosphere-free combination.
2. Tropospheric correction.
3. Satellite clock offset from GPS time.

In summary, an estimate using the least squares technique (LS) and an estimate using the ridge-type technique (RT) was made for three different observation cases. Several systems were solved for each case.

5. Results

The reduction of the condition number was evaluated in terms of the number of equations. The figures 1-2 show the reduction for two data sets, D1 and D10. The data set D5 is the same that D1 but considering different number of equations, the results are very similar to that of figure 1. More detailed comparison graphs can be obtained from the author. We can observe a very erratic behaviour at the highest equation numbers for both matrix models. Figures 3-4 show solutions dx and dy for both methods considering data set D1. Similar trends hold for variables dz and dt. Figure 5 shows how the bias introduces an increase in the root mean square (RMS) for the RT method. In order to evaluate the relative efectiveness of the solution methods we have used the relative error. In figure 6 it should be noted the improvement in this error for RT method. The results with the D10 data set are similar to the previous obtained results, see figures 7, 8 ,11 and 12. However the increase in value of RMS becomes more pronounced as the number of equations increase (figure 9) . Moreover the inprovement in the solution is better when the number of equations increase (figure 10).

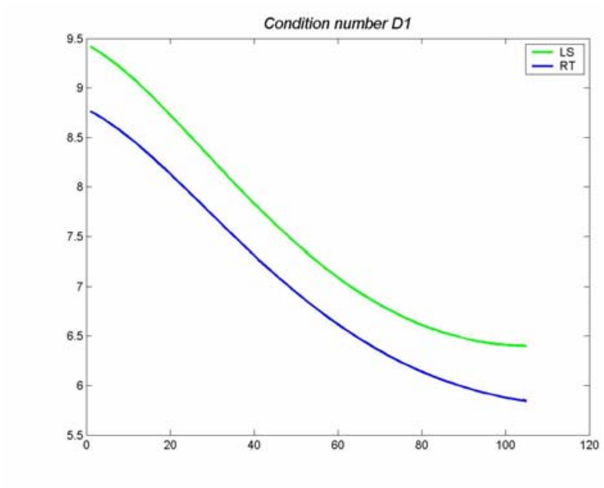

figure 1: condition number D1

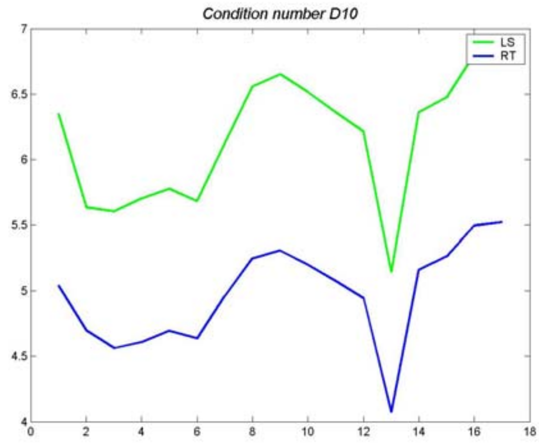

figure 2: condition number D10

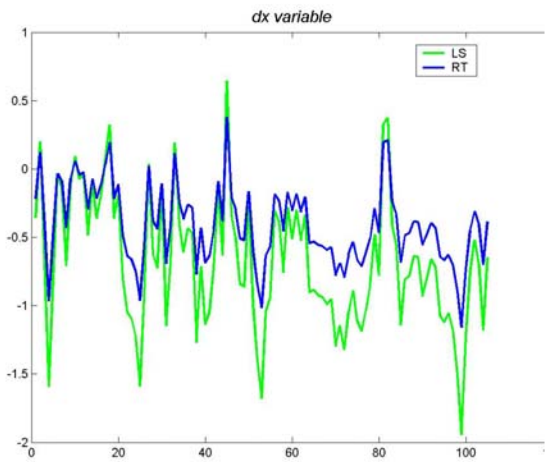

figure 3: Compared solutions dx LS and RT (D1)

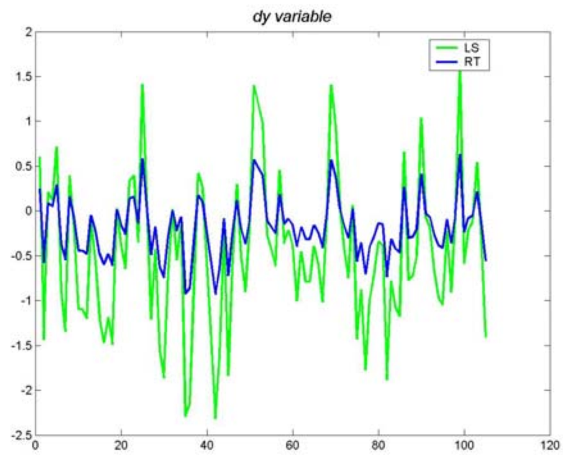

figure 4: Compared solutions dy LS and RT (D1)

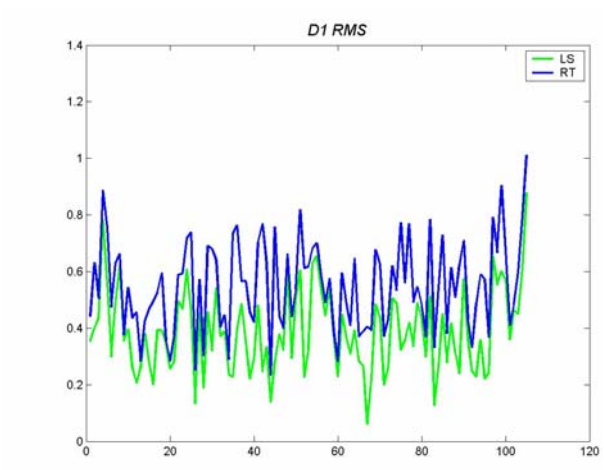

figure 5: RMS D1

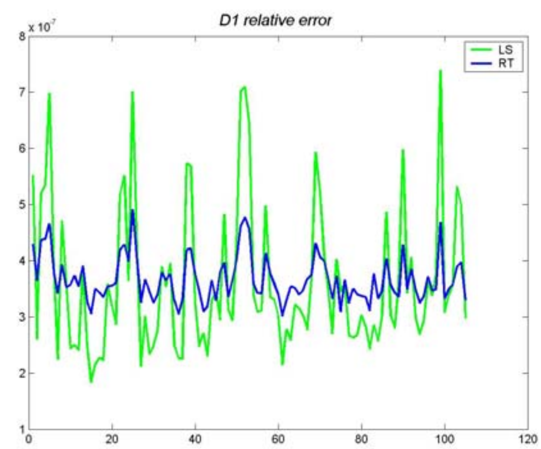

figure 6: Relative error D1

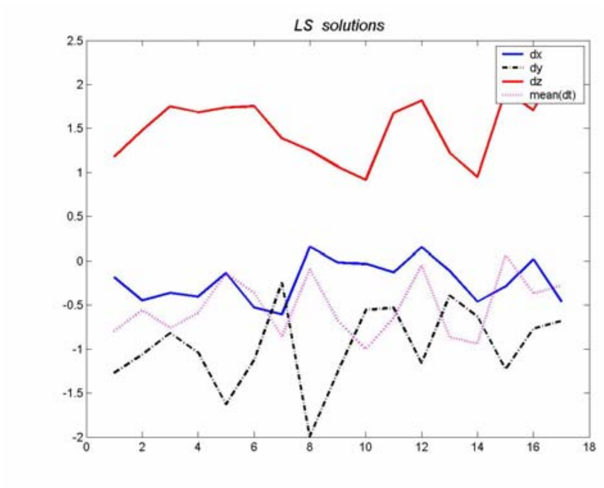

figure 7: Solutions LS (D10)

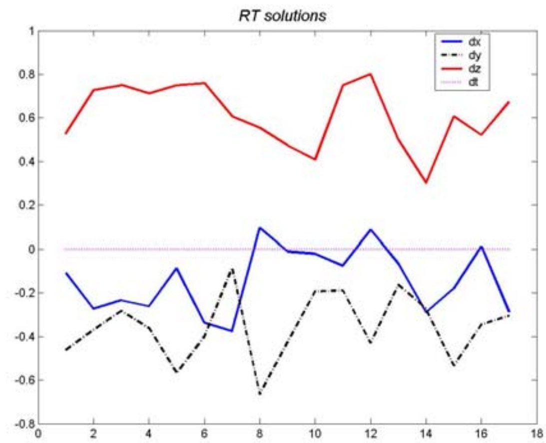

figure 8: Solutions RT (D10)

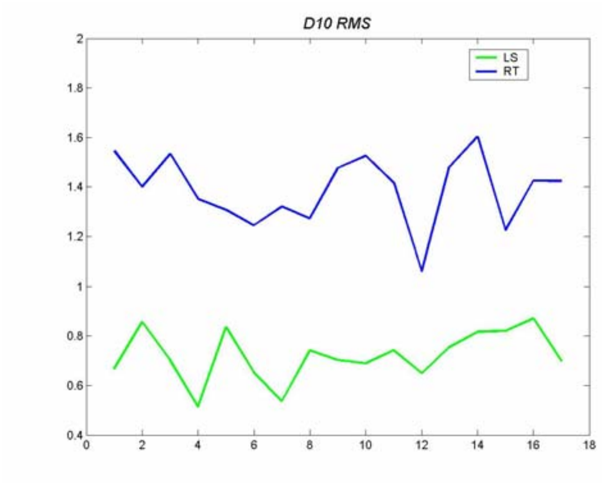

figure 9: RMS D10

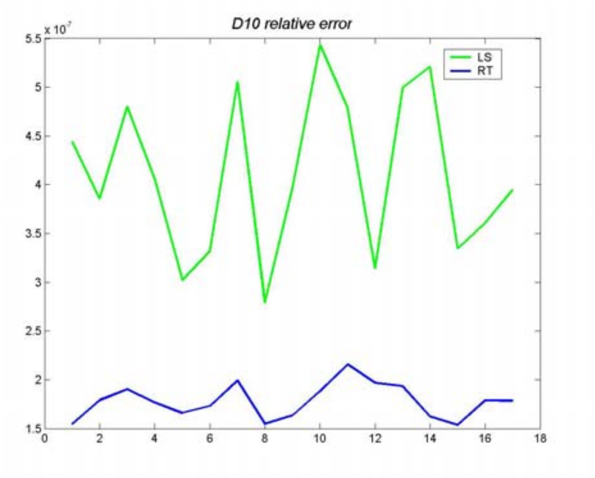

figure 10: Relative error D10

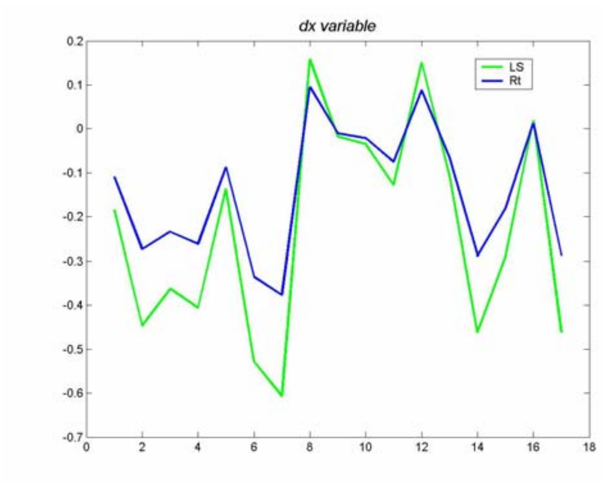

figure 11: Compared solutions dx LS and RT (D10)

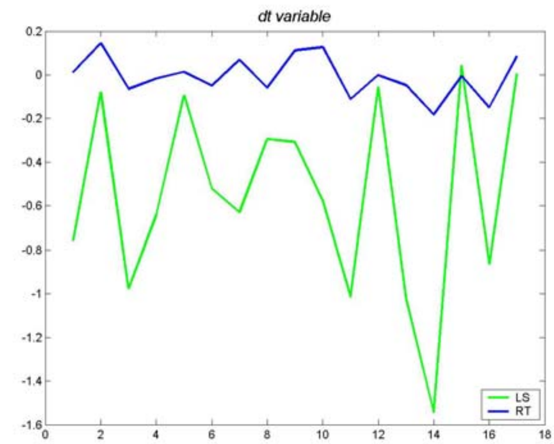

figure 12: Compared solutions dt LS and RT (D10)

References

[1] D.W. Marquardt, An algorithm for least-squares estimation of non-linear parameters, J. Society of industrial an Appl. Math. 11 (2) (1963) 431-441.

[2] D.A. Cicci, Optimal A priori Covariance Selection for the solution of Ill-Conditioned Nonlinear Inverse Problems, Doctoral dissertation, The University of Texas at Austin, Department of Aerospace Engineering and Engineering Mechanics, August 1987.

[3] G. Strang, Linear Algebra and Its Applications, 2nd (ed.), Academic Press, New York, 1980.

[4] G. Strang, K. Borre, Linear algebra, geodesy, and GPS, Wellesley-Cambridge Press, 1997.

[5] S.D. Silvey, Multicollinearity and imprecise estimation, Technometrics 11 (1969) 539-552.