GPS, INTEGERS, ADJUSTMENT AND PROBABILITY

Peter J.G. Teunissen

ABSTRACT

In this contribution we give a brief review of the theory of integer estimation as it has been developed for use with the Global Positioning System (GPS). First we consider the GPS observation equations (*GPS and Integers*). These observation equations are not of the usual type, since some of the parameters are known to be integer. This implies that an extention of 'classical' adjustment theory is needed. Such an extention is presented by showing ways of how to solve integer adjustment problems (*Integers and Adjustment*). Different integer estimators are given and a class of admissible integer estimators is defined. Next we consider some qualitative aspects of an integer adjustment (*Adjustment and Probability*). It is argued that the usual qualitative description by means of second moments or variance-covariance matrices is not sufficient. A direct probabilistic description is needed instead. Such a description is presented by means of the probability mass function of the integer ambiguities. For GPS ambiguity resolution, the probability of correct integer estimation is particularly of interest (*Probability and GPS*). It describes the ambiguity succes rate and shows whether or not the estimated ambiguities may be treated deterministically. Since different integer estimators have different success rates, one is particularly interested in the estimator which maximizes this success rate. The answer is given by the theorem provided.

1 GPS and Integers

As witnessed by the enormous GPS literature available, there exists a great variety in GPS models currently in use. They may range from single- site models used for local monitoring purposes to multi-baseline models used as a tool for studying geodynamic phenomena, or from supershort multi-baseline models used for local attitude determination to wide-area models for transmitting differential corrections. Depending on the application at hand, each one of these models may differ in the way the observed signals are linked, in the way the reference systems and the orbits are treated, or in the way the receiver and propagation delays are modelled. An overview of these and other GPS models, together with their applications in surveying, navigation, geodesy and geophysics, can be found in textbooks such as [*Hofmann-Wellenhof et al.*, 1997], [*Leick*, 1995], [*Parkinson and Spilker*, 1997], [*Strang and Borre*, 1997] and [*Teunissen and Kleusberg*, 1998].

Despite the differences in application of the various GPS models, their observation equations are to a large part the same. For the single baseline case, using two receivers, each tracking the same two satellites, the double differenced (DD) phase and code observation equations for a single epoch i read

$$\begin{array}{rcl} \phi_1(i) & = & \rho(i) - \mu_1 I(i) + T(i) + \lambda_1 a_1 \\ \phi_2(i) & = & \rho(i) - \mu_2 I(i) + T(i) + \lambda_2 a_2 \\ p_1(i) & = & \rho(i) + \mu_1 I(i) + T(i) \\ p_2(i) & = & \rho(i) + \mu_2 I(i) + T(i) \end{array} \quad (1)$$

where ϕ_1 and ϕ_2 are the DD phase observables on L_1 and L_2; p_1 and p_2 are the DD code

observables on L_1 and L_2; ρ is the DD-form of the unknown receiver-satellite range; I and T are the DD-forms of resp. the unknown ionospheric and troposheric delay, and a_1 and a_2 are the unknown integer ambiguities. The known wavelengths are denoted as λ_1, λ_2. Since the ionospheric delay is to a first order inversely proportional to the square of the frequency, we have to the same degree of approximation $\mu_1 = \frac{\lambda_1}{\lambda_2} = \frac{60}{77}$, $\mu_2 = \frac{\lambda_2}{\lambda_1} = \frac{77}{60}$.
Depending on the application the atmospheric delays may be assumed present or absent. For sufficiently short baselines, the ionospheric delays can usually be assumed absent. The same holds true for the troposheric delays in case the differences in height are not too large. When one is forced to assume the delays to be present, a strenghtening of the model is sometimes still possible. This can be done by either using an *a priori* weighting of the delays or by using a further parametrization in fewer parameters. For instance, since the ionosphere decorrelates spatially, the ionospheric delays are often weighted a priori as function of the baseline length. And in case of the troposphere, a parametrization in fewer parameters is often done using the so-called mapping functions.
Although the above equations are already useful for many applications, they are not yet useful for positioning purposes. For that to be the case, a further parametrization of $\rho(i)$ into the unknown baseline vector between the two receivers is needed. Such models are referred to as *geometry-based*, because of the explicit presence in the model of the relative receiver-satellite geometry. Without such a parametrization the model is referred to as *geometry-free*.
In all applications where the above equations form the backbone of the particular GPS model used, the unknown carrier phase ambiguities a_i, a_2 are known to be *integer*. Hence, despite the differences in application of the various GPS models, there should in principle be no difference in how these integer parameters are dealt with. Their ambiguity resolution problems should intrinsically be the same. This implies that it should be possible to develop a single theoretical framework that applies to every GPS model for which ambiguity resolution would make sense. Such a framework did not exist in the pre-GPS era, simply because 'classical' adjustment theory is not equiped to deal with integer parameters. Even in the GPS era it took quite some time before the theoretical framework, as we know it today, took its shape. It started with the development of methods of integer estimation. Although some of the methods proposed in the beginning were ad hoc, inefficient and sometimes even simply wrong, the current state of affairs is that one can indeed speak of a reasonable mature theory of integer estimation. However, this can not yet be said of the necessary probabilistic theory. It is amazing to see that in almost all contributions the probabilistic part is either ignored completely or thought to be solved by 'classical' means such as those of 'the linear model'. It is only recently that integer adjustment has been complemented to some extent with some of the necessary probabilistic theory. In this contribution we will confine ourselves to this probabilistic part and briefly review some of the results as obtained in [*Teunissen*, 1997, 1998, 1999].

2 Integers and Adjustment

In principle all GPS models that are based on (1) can be cast in the following conceptual frame of linear(ized) observation equations

$$y = Aa + Bb + e \tag{2}$$

where y is the given GPS data vector of order m, a and b are the unknown parameter vectors respectively of order n and o, and where e is the noise vector. The matrices A and B are the corresponding design matrices. They are assumed to be of full rank. The data vector y will usually consist of the 'observed minus computed' single- or dual- frequency double-difference (DD) phase and/or pseudorange (code) observations accumulated over all observation epochs. The entries of vector a are then the DD carrier phase ambiguities, expressed in units of cycles rather than range. They are known to be *integers*, $a \in Z^n$. The entries of the vector b will consist of the remaining unknown parameters, such as for instance receiver-satellite ranges in case of the geometry-free model or baseline coordinates in case of the geometry-based model, and possibly atmospheric delay parameters as needed for the troposphere and/or ionosphere. All the parameters collected in b are real-valued.

The procedure which is usually followed for solving the GPS model (2), can be divided into three steps. In the *first* step one simply disregards the integer constraints $a \in Z^n$ on the ambiguities and performs a standard adjustment. As a result one obtains the (real-valued) estimates of a and b, together with their variance-covariance (vc-) matrix

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}, \begin{bmatrix} Q_{\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}} \end{bmatrix} \tag{3}$$

In GPS terminology this solution is referred to as the 'float' solution. In the *second* step the 'float' ambiguity estimate $\hat{a}$ is used to compute the corresponding integer ambiguity estimate $\check{a}$. This implies that a mapping $F : R^n \mapsto Z^n$ is introduced, from the n-dimensional space of reals to the n-dimensional space of integers, such that

$$\check{a} = F(\hat{a}) \tag{4}$$

Once these integer ambiguities are computed, they are used in the *third* step to finally correct the 'float' estimate of b. As a result one obtains the 'fixed' solution $\check{b} = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1}(\hat{a} - \check{a})$.

It will be clear that the actual integer adjustment is confined to the second step of above, in particular to the choice of the map F. Due to the discrete nature of Z^n, the map F will not be one-to-one, but instead a many-to-one map. This implies that different real-valued ambiguity vectors will be mapped to the same integer vector. One can therefore assign a subset $S_z \subset R^n$ to each integer vector $z \in Z^n$:

$$S_z = \{x \in R^n \mid z = F(x)\}, \quad z \in Z^n \tag{5}$$

The subset S_z contains all real-valued ambiguity vectors that will be mapped by F to the same integer vector $z \in Z^n$. This subset is referred to as the *pull-in-region* of z. It is the region in which all ambiguity 'float' solutions are pulled to the same 'fixed' ambiguity vector z.

Since the pull-in-regions define the integer estimator completely, one can define classes of integer estimators by imposing various conditions on the pull-in-regions. Such a class was introduced by the author and referred to as the class of admissible integer estimators. These integer estimators are defined as follows:

Definition

An integer estimator is said to be *admissible* if

$$\begin{array}{ll} (i) & \bigcup_{z \in Z^n} S_z = R^n \\ (ii) & S_{z_1} \bigcap S_{z_2} = \{0\}, \ \ \forall z_1, z_2 \in Z^n, z_1 \neq z_2 \\ (iii) & S_z = z + S_0, \ \ \forall z \in Z^n \end{array} \tag{6}$$

This definition is motivated as follows. Each one of the above three conditions describe a property of which it seems reasonable that it is possessed by an arbitrary integer ambiguity estimator. The first condition states that the pull-in-regions should not leave any gaps and the second that they should not overlap. The absence of gaps is needed in order to be able to map any 'float' solution $\hat{a} \in R^n$ to Z^n, while the absence of overlaps is needed to guarantee that the 'float' solution is mapped to just one integer vector. Note that we allow the pull-in-regions to have common boundaries. This is permitted if we assume to have zero probability that $\hat{a}$ lies on one of the boundaries. This will be the case when the probability density function (pdf) of $\hat{a}$ is continuous.

The third and last condition follows from the requirement that $F(x+z) = F(x)+z, \forall x \in R^n, z \in Z^n$. Also this condition is a reasonable one to ask for. It states that when the 'float' solution is perturbed by $z \in Z^n$, the corresponding integer solution is perturbed by the same amount. This property allows one to apply the *integer remove-restore* technique: $F(\hat{a} - z) + z = F(\hat{a})$. It therefore allows one to work with the fractional parts of the entries of $\hat{a}$, instead of with its complete entries.

There exist many admissible integer estimators. In fact, one can 'invent' one's own admissible integer estimator by simply specifying pull-in-regions that satisfy the above definition. Here we will give as an example the pull-in-regions of three more commonly used admissible estimators. The simplest integer map is the one that corresponds to an integer rounding. In this case the integer vector is obtained from a rounding of each of the entries of $\hat{a}$ to its nearest integer. Since componentwise rounding implies that each real-valued ambiguity estimate $\hat{a}_i, i = 1, \ldots, n$, is mapped to its nearest integer, the absolute value of the difference between the two is at most $\frac{1}{2}$. The subsets $S_{R,z}$ that belong to this integer estimator are therefore given as

$$S_{R,z} = \cap_{i=1}^{n} \{\hat{a} \in R^n \mid \mid \hat{a}_i - z_i \mid \leq \frac{1}{2}\} \ , \ \forall z \in Z^n \tag{7}$$

The subset $S_{R,z}$ is an n-dimensional cube, with sides of length 1 and centred at the grid point z.

Another relatively simple integer ambiguity estimator is the bootstrapped estimator. This estimator can be seen as a generalization of the previous one. It still makes use of integer rounding, but it also takes some of the correlation between the ambiguities into account. The bootstrapped estimator results from a sequential conditional least- squares adjustment and is computed as follows. If n ambiguities are available, one starts with the first ambiguity $\hat{a}_1$, and rounds its value to the nearest integer. Having obtained the integer value of this first ambiguity, the real-valued estimates of all remaining ambiguities are then corrected on the basis of their correlation with the first ambiguity. Subsequently the second, but now corrected, real-valued ambiguity estimate is rounded to its nearest integer. Having obtained the integer value of the second ambiguity, the real-valued estimates of all remaining $n-2$ ambiguities are again corrected, but now on the basis of their correlation with the second ambiguity. This process of rounding and correcting is continued until all ambiguities are taken care of.

With c_i denoting the ith canonical unit vector having a 1 as its ith entry, the pull-in-regions $S_{B,z}$ that belong to the bootstrapped estimator can be shown to be given as

$$S_{B,z} = \cap_{i=1}^{n} \{\hat{a} \in R^n \mid \mid c_i^T L^{-1}(\hat{a} - z) \mid \leq \frac{1}{2}\} \ , \ \forall z \in Z^n \tag{8}$$

with matrix L being the lower triangular unit matrix that follows from applying a triangular decomposition to the variance-covariance matrix of $\hat{a}$. Note that these pull-in-regions reduce to the ones of (7) when L becomes diagonal. This is the case when the ambiguity variance-covariance matrix is diagonal. In that case the two integer estimators $\check{a}_R$ and $\check{a}_B$ are identical. The third admissible estimator of which the pull-in-region will be given is the integer least-squares estimator. By again using the LDU-decomposition of $Q_{\hat{a}}$ the least-squares' pull-in-region reads

$$S_{LS,z} = \cap_{c_i \in L^{-1}(Z^n)} \{\hat{a} \in R^n \mid \ \mid c_i^T D^{-1} L^{-1}(\hat{a} - z) \mid \leq \frac{1}{2} c_i^T D^{-1} c_i\} \tag{9}$$

Note that (9) and (8) become identical when the matrix entries of L^{-1} are all integer. This is the case when L is an admissible ambiguity transformation.

3 Adjustment and Probability

In the previous section it was shown how an integer adjustment can be performed. One first needs to define the type of integer estimator. This is done by choosing the integer map $F : R^n \leftrightarrow Z^n$

or by choosing the corresponding pull-in-regions $S_z \subset R^n$. Once the type of integer estimator is chosen, the actual adjustment can be executed. This is done by searching for the pull-in-region in which the data vector lies. The integer solution follows then once this pull-in-region is identified.
It will be clear that a mere adjustment (or parameter estimation) is not enough. One also needs to be able to describe the quality of the adjustment result. After all one can always perform an adjustment whether the data are of good quality or not. In 'classical' adjustment theory a special place is reserved for the variance-covariance (vc-) matrix. It is often the vc-matrix of the estimated parameters which is used to decribe the quality. In 'classical' adjustment theory, the vc-matrix is even used as criterion for deriving optimal estimators. The least-squares estimator for instance, is known to be optimal in the sense that its variance is the smallest of all linear unbiased estimators.
Although the vc-matrix is often used in practice, the reason for its usage as quality measure is not always stated explicitly and unequivocally. Why is the vc-matrix used and what does it tell us? The vc-matrix of a random vector is defined as the second (central) moment of the vector's distribution. It describes how large the squared differences between sample values and the mean (the first moment) will be *on average*. Thus the vc-matrix is a measure of the expected spread around the random vector's mean. In practice the term 'precision' is used for this notion. But although such precision information is of importance in its own right, it is not the only reason for the popular usage of the vc-matrix. In practice, the popularity of the vc-matrix also stems from the fact that often the vc-matrix is used as a tool for constructing confidence regions. Standard ellipses for instance, are often interpreted as such confidence regions. This particular usage of the vc-matrix implies however that one assumes the random vector to be normally distributed. Since the normal distribution is completely specified once its first two moments are known, knowledge of the vc-matrix is sufficient for determining confidence regions and for determining the probability that the outcome of the estimator stays within a certain limit from its mean. In general however, the information content of the vc-matrix is not enough to determine confidence regions. For making such probability statements, the complete probability distribution is needed and not only its second (central) moment. After all, different distributions can have identical vc-matrices. The definition of the second (central) moment is not even restricted to continuous, unimodel distributions, but applies to discrete and multimodal distributions as well.
The above makes clear that in our case of an integer adjustment, the concept of the vc-matrix is much less useful. Although the vc-matrix still describes the 'expected spread', it can not be used anymore to determine probabilities and confidence regions. After all, in case of an integer adjustment the distribution of the estimator will be of the discrete type, thus not continuous and certainly not normal or Gaussian. The distribution of the integer estimator $\check{a}$ will be a probability *mass* function (pmf), which we shall denote as $P(\check{a} = z)$, with $z \in Z^n$. Thus $P(\check{a} = z)$ denotes the probability that the ambiguity vector $\check{a}$ equals the integer vector $z \in Z^n$. In order to determine this distribution, we first need the probability density function (pdf) of $\hat{a}$. The pdf of $\hat{a}$ will be denoted as $p_a(x)$, with $x \in R^n$. The subindex is used to show that the pdf still depends on the unknown parameter vector $a \in Z^n$. In case the pdf of $\hat{a} \in R^n$ is *elliptically contoured* it is of the form

$$p_a(x) = \sqrt{\det(Q_{\hat{a}}^{-1})} G(\| x - a \|_{Q_{\hat{a}}}^2) \qquad (10)$$

where $G : R \mapsto [0, \infty)$ is decreasing and $Q_{\hat{a}}$ is positive-definite. Several important distributions belong to this family. The multivariate normal distribution can be shown to be a member of this family by choosing $G(x) = (2\pi)^{-\frac{n}{2}} \exp -\frac{1}{2}x$, $x \in R$. Another member is the multivariate t-distribution.
The pmf of $\check{a}$ can now be obtained as follows. Using the concept of the pull-in-region, the integer estimator is defined as $\check{a} = z \iff \hat{a} \in S_z$. This shows that $P(\check{a} = z) = P(\hat{a} \in S_z)$. The probability that $\check{a}$ coincides with z is therefore given by the integral of the pdf $p_a(x)$ over the

pull-in-region $S_z \subset R^n$. Hence, the pmf of $\check{a}$ follows as

$$P(\check{a} = z) = \int_{S_z} p_a(x)dx \ , \ \forall z \in Z^n \tag{11}$$

It is this function which gives a complete description of the random characteristics of the integer ambiguity estimator $\check{a}$. Hence, it is this function which should be used when studying the qualitative aspects of the integer estimator.

4 Probability and GPS

The quality of the integer ambiguity estimator is particularly of interest in case of GPS. In case of GPS one usually treats the estimated integer ambiguities as if they were deterministic variates. Theoretically this is not correct as the previous section has shown. Neglecting the random nature of the estimated integer ambiguities when applying the error propagation law to

$$\check{b} = \hat{b} - Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} (\hat{a} - \check{a}) \tag{12}$$

implies that a too optimistic quality description is obtained for the so-called 'fixed' estimator $\check{b}$. Whether this is acceptable or not depends of course on the approximation involved. One should therefore have a diagnostic tool available on the basis of which one can decide whether the approximation is acceptable or not. This diagnostic tool is provided by the *ambiguity success rate*, which is defined as the probability of correct integer estimation

$$\text{ambiguity success rate} = P(\check{a} = a) \tag{13}$$

One should therefore first compute the ambiguity success rate and check whether it is sufficiently close to one, before deciding that a deterministic treatment of $\check{a}$ is acceptable.
Note that the pmf (11) as well as the success rate (13) still depend on the type of pull-in-region and thus on the type of integer estimator chosen. Changing the geometry of the pull-in-region will change both the pmf and the ambiguity success rate. It is therefore not only of theoretical interest, but also of practical interest, to know which integer estimator maximizes the ambiguity success rate. The answer is given by the following theorem:

Theorem (*Teunissen*)
Let the pdf of $\hat{a}$ be elliptically contoured and the integer least-squares estimator be given as

$$\check{a}_{LS} = \arg \min_{z \in Z^n} \parallel \hat{a} - z \parallel^2_{Q_{\hat{a}}}$$

Then

$$P(\check{a}_{LS} = a) \geq P(\check{a} = a) \tag{14}$$

for any admissible estimator $\check{a}$.

This theorem gives a probabilistic justification for using the *integer* least-squares estimator. As a probabilistic justification it may be considered to replace the theorem of Gauss which states that the *real-valued* least-squares estimator has smallest variance of all linear unbiased estimators. The theorem particularly applies to GPS ambiguity resolution, for which often the multivariate normal distribution is assumed to hold true. For GPS ambiguity resolution one is thus better off using the integer least-squares estimator than any other admissible integer estimator, such as, for instance, the 'rounding' estimator or 'bootstrapped' estimator.

5 References

[1] Hofmann-Wellenhof, B., H. Lichtenegger, J. Collins (1997): *Global Positioning System: Theory and Practice*. 4th edition. Springer Verlag.

[2] Leick, A. (1995): *GPS Satellite Surveying*, 2nd edition, John Wiley and Sons, New York.

[3] Parkinson, B., J.J. Spilker (eds) (1996): *GPS: Theory and Applications*, Vols 1 and 2, AIAA, Washington DC.

[4] Strang, G. and K. Borre (1997): *Linear Algebra, Geodesy, and GPS*, Wellesley-Cambridge Press.

[5] Teunissen, P.J.G. (1997): Some remarks on GPS ambiguity resolution. *Artificial Satellites*, Vol. 32, No. 2, pp. 119-130.

[6] Teunissen, P.J.G. (1998): On the integer normal distribution of the GPS ambiguities. *Artificial Satellites*, Vol. 33, No. 2, pp. 49-64.

[7] Teunissen, P.J.G. (1999): A theorem on maximizing the probability of correct integer estimation. *Artificial Satellites*, Vol. 34, No. 2.

[8] Teunissen, P.J.G., A. Kleusberg (eds) (1998): *GPS for Geodesy*, 2nd enlarged edition, Springer Verlag.