OPTIMIZATION OF LEAST-SQUARES COMPUTATIONS IN ON-LINE PHOTOGRAMMETRY

Dr. J.A.R. Blais
Division of Surveying Engineering
The University of Calgary
Calgary, Alberta, Canada T2N 1N4
Commission III

Abstract:

Least-squares computations can be carried out in several different ways for adjustment, prediction, filtering and smoothing purposes. The orthogonal and square-root algorithms are especially advantageous for on-line computations where computational efficiency, storage requirements and numerical accuracy are critical. Different algorithms for such least-squares computations are briefly discussed with their principal characteristics in interactive environments. Some aspects of the corresponding firmware implementations are also mentioned for photogrammetric applications.

1. INTRODUCTION

Over the past decade, there has been much research and development in the field of least-squares computations for adjustment, prediction, filtering and smoothing purposes. A great deal of motivation for these efforts has been related to complex navigation problems in the space programme. The critical situations often encountered in real-time navigation require the best numerical techniques available.

The estimation problems in the context of on-line photogrammetry are obviously much simpler than in space navigation. However in order to optimize the benefits of recursive least-squares computations with limited computer facilities and quality control requirements, it is imperative that proper attention be given to the computational algorithms and related numerical matters. Otherwise the advantages of on-line computations can easily be eroded away to the point that their implementation becomes debatable in practical situations.

The principal objective of the following discussion is to emphasize that least-squares computations can indeed be optimized in terms of computational efficiency, storage requirements and numerical accuracy without sacrificing any aspect of the mathematical rigour of the least-squares formulations.

2. LEAST-SQUARES ESTIMATION PROBLEMS

Photogrammetric estimation problems are quite varied in different application contexts. For the purposes of the following discussion, they are going to be classified as adjustment, prediction, filtering and smoothing problems of least-squares estimation. Appropriate linearizations of the mathematical models are also assumed for the sequel.

Adjustment problems for orientation and triangulation purposes are definitely the most common least-squares problems in photogrammetry. The corresponding mathematical model consists in forming an overdetermined system of linear equations

$$Ax = f + e \qquad E[e] = 0 \qquad E[e\,e^T] = C_e$$

with a design matrix A, unknown parameters x, right-hand-side observation vector f and observational errors with zero expectation and covariance matrix C_e. The well-known least-squares solution is simply written as

$$\hat{x} = (A^T C_e^{-1} A)^{-1} A^T C_e^{-1} f$$

with corresponding covariance matrix

$$C_{\hat{x}} = (A^T C_e^{-1} A)^{-1}$$

for the estimated parameters.

Prediction problems for error correction and signal estimation purposes are also common least-squares problems in photogrammetry. Their formulation is simply the dual of the preceding situation with an underdetermined system of linear equations

$$Ax = f + e \qquad E[x] = 0 \qquad E[x x^T] = C_x$$

with a design matrix A, unknown parameters x with zero expectation and covariance matrix C_x, and a right-hand-side observation vector f. The well-known least-squares solution is simply written as

$$\hat{x} = C_x A^T (A C_x A^T)^{-1} f$$

with corresponding covariance matrix

$$C_{\hat{x}} = C_x - C_x A^T (A C_x A^T)^{-1} A C_x$$

for the estimated parameters.

Before considering other least-squares estimation problems, it is important to note that the previous two situations can be combined into a single formulation in terms of the linear equations:

$$Ax = f + e \qquad E[e] = 0 \qquad E[e e^T] = C_e$$
$$E[x] = 0 \qquad E[x x^T] = C_x$$

using the same notation as before. The general least-squares solution is then written as

$$\hat{x} = (A^T C_e^{-1} A + C_x^{-1})^{-1} A^T C_e^{-1} f$$
$$= C_x A^T (A C_x A^T + C_e)^{-1} f$$

with corresponding covariance matrix

$$C_{\hat{x}} = (A^T C_e^{-1} A + C_x^{-1})^{-1}$$
$$= C_x - C_x A^T (A C_x A^T + C_e)^{-1} A C_x$$

for the estimated parameters, assuming the nonsingularity of the matrices which are to be inverted. The first expressions for $\hat{x}$ and $C_{\hat{x}}$ are in terms of the inverses of the covariance matrices (i.e. statistical weight or information matrices) while the second are in terms of covariance matrices.

Filtering and smoothing problems arise whenever a system process is considered in conjunction with an observation process. For instance with a servo-driven photogrammetric instrument, it would seem appropriate to model the dynamics of the instrument between observations as well as the observations themselves. Therefore in addition to the previously defined observation process

$$A_k x_k = f_k + e_k \qquad E[e_k] = 0 \qquad E[e_k e_k^T] = C_{e_k}$$

for the k-th step, let us consider a linear system process modelled as

$$x_{k+1} = S_k x_k + u_k \qquad E[u_k] = 0 \qquad E[u_k u_k^T] = C_{u_k}$$

for the transition from the k-th to the (k+1)th step, in which S_k denotes the transition matrix. The estimate of state vector x_k is then updated in two operations per step:

$$\hat{\hat{x}}_{k-1} \longrightarrow \hat{x}_k \longrightarrow \hat{\hat{x}}_k \qquad k = 1, 2, \ldots$$

corresponding to the system and observation processes. Explicitly, for the system process,

$$\hat{x}_k = S_{k-1} \hat{\hat{x}}_{k-1}$$

with the corresponding updating of the covariance matrix

$$C_{\hat{x}_k} = S_{k-1} C_{\hat{\hat{x}}_{k-1}} S_{k-1}^T + C_{u_{k-1}}$$

For the observation process,

$$\hat{\hat{x}}_k = \hat{x}_k + (A_k^T C_{e_k}^{-1} A_k + C_{\hat{x}_k}^{-1})^{-1} A_k^T C_{e_k}^{-1} (f_k - A_k \hat{x}_k)$$

$$= \hat{x}_k + C_{\hat{x}_k} A_k^T (A_k C_{\hat{x}_k} A_k^T + C_{e_k})^{-1} (f_k - A_k \hat{x}_k)$$

with the corresponding updating of the covariance matrix

$$C_{\hat{\hat{x}}_k} = (A_k^T C_{e_k}^{-1} A_k + C_{\hat{x}_k}^{-1})^{-1}$$

$$= C_{\hat{x}_k} - C_{\hat{x}_k} A_k^T (A_k C_{\hat{x}_k} A_k^T + C_{e_k})^{-1} A_k C_{\hat{x}_k}$$

assuming the nonsingularity of the matrices which are to be inverted. Again, the first expressions are in terms of statistical weight or information matrices while the second are in terms of covariance matrices.

The smoothing of the state vectors simply corresponds to the weighted average of the forward and backward filtered state vectors for each step [e.g., Gelb, 1974]. Algebraically, the computations can be carried out recursively using exactly the same algebra as for filtering with appropriate substitutions for the state vector and covariance matrix estimates [Blais, 1984]. Hence for optimization purposes, filtering and smoothing present exactly the same challenges in terms of computational efficiency, storage requirements and numerical accuracy. In the following discussion, smoothing problems need not be analyzed separately from filtering problems at least for the photogrammetry application context.

3. ORTHOGONAL TRANSFORMATION METHODS

The algebraic formulations of the preceding section are often called covariance and inverse covariance (or information) formulations of least-squares estimation. Such formulations have been known to cause numerical stability problems in certain contexts. Essentially, the numerical evaluation of the covariance or information matrix has resulted in negative eigenvalues in some instances and diverging covariance filtering in other cases [Chin, 1983].

Fortunately, other approaches are available to optimize the numerical computations and avoid ill-conditioning in most cases. Orthogonal transformation methods provide such an approach that has led to several square-root and other algorithms over the past ten years for a wide spectrum of applications. In terms of numerical conditioning, the condition number associated with a square-root method is approximately the square root of the corresponding condition number for a covariance or information method with the same observation and/or system data [Kaminski et al, 1971].

Among the best known orthogonal methods are the following three:
(a) the Gram-Schmidt transformations which are based on projections,
(b) the Householder transformations which are based on reflections,
(c) the Givens transformations which are based on rotations.
The latter method is especially known for its advantageous characteristics in terms of data storage requirements, numerical stability and computational efficiency [Gentleman, 1973; Lawson and Hanson, 1974; George and Heath, 1980; Blais, 1982 and 1983].

Explicitly, considering the simplest least-squares adjustment problem that implies the following overdetermined system of m linear equations with n unknowns, where $m \geq n$,

$$Ax = f + e \qquad E[e] = 0 \qquad E[e\,e^T] = C_e = I$$

with the same notation as before, the corresponding normal equations are

$$A^T A x = A^T f$$

using unit statistical weights. The usual solution approach consists in using Choleski's square-root algorithm to transform this linear system into an upper triangular one $Ux = d$ which can then be solved easily by back substitution.

Alternatively, if an orthogonal decomposition $A = QR$ is available, where Q is a $m \times n$ matrix whose columns are orthogonal and R is a $n \times n$ upper triangular matrix, then the normal equations can be rewritten as

$$\begin{aligned} A^T A x &= R^T Q^T Q R x \\ &= R^T R x \\ &= R^T Q^T f \end{aligned}$$

and since R is nonsingular whenever $A^T A$ is, then

$$R x = Q^T f$$

These reduced normal equations are the same as the previous ones, that is, $Ux = d$, as $A^T A = R^T R$ is actually the Choleski factorization carried out with the square-root algorithm. It then follows that $U = R$ and $d = Q^T f$.

A particularly advantageous approach for the evaluation of U and d is with the application of a sequence of Givens transformations to A and f, without any explicit requirement of the Q matrix. These Givens transformations are simply two-dimensional rotations which permit a row-by-row transformation of A and f into U and d, respectively. This implies that in the usual context, the observations can be processed sequentially into an upper triangular system of normal equations. The corresponding solution can obviously be obtained at any stage of the data processing by simply carrying out a back substitution into the upper triangular system $Ux = d$.

The actual definitions of those Givens rotations necessary for some arbitrary observation vector and the corresponding updating of U and d are given in details in [Blais, 1982 and 1983; Teskey, 1983]. In the current context of on-line photogrammetric computations, one is not only interested in the parameter estimate $\hat{x}$ but also in the corresponding error estimate $C_{\hat{x}}$, the covariance matrix for $\hat{x}$. Recalling that

$$C_{\hat{x}} = (A^T A)^{-1} = (U^T U)^{-1} = U^{-1} U^{-T}$$

with $C_e = I$ as before, it becomes greatly advantageous to update U^{-1} simultaneously with U and d for each observation. This can be done in a straightforward manner using the Givens transformations to update U and d, and the transposed inverse of U for each observation. The details of these computations for arbitrary C_e are again given in [Blais, 1983; Teskey, 1983].

This approach to recursive least-squares computations can also be easily extended to the filtering context for a state vector x whose estimate is updated in two operations per step:

$$\hat{\hat{x}}_{k-1} \longrightarrow \hat{x}_k \longrightarrow \hat{\hat{x}}_k \qquad k = 1, 2, \ldots$$

corresponding to the system and observation processes, respectively. The preceding discussion of the updating of U, U^{-T} and d for each observation using Givens transformations applies also to the operations $\hat{x}_k \longrightarrow \hat{\hat{x}}_k$ with any number of observations. The remaining prediction part of the filtering procedure, that is, $\hat{\hat{x}}_{k-1} \longrightarrow \hat{x}_k$, is quite straightforward to formulate using the transition matrix S_{k-1} and the corresponding covariance matrix $C_{u_{k-1}}$. The details of these computations for filtering and hence for smoothing are given in [Blais, 1984].

4. OTHER COMPUTATIONAL APPROACHES

The preceding Givens transformation method has also been modified to avoid all square-root operations and hence be significantly more efficient [Gentleman, 1973]. However, in some contexts, numerical stability problems can arise unless some close numerical monitoring is done on the computations. In other words, the advantageous numerical stability characteristics of the Givens transformations can easily be eroded away if their mechanization is not carried out with proper care.

Another formulation which is closely related to the square-root-free Givens transformations with scalar observations is the so-called U-D factorization approach, where U and D are the upper triangular and diagonal factors, respectively [Agee and Turner, 1972; Bierman, 1977]. The restriction to scalar observations is no problem as any vector observation can be treated as a sequence of scalar observations.

Other formulations have also been developed in various application contexts with different measures of success in terms of computational efficiency, numerical stability and storage requirements [e.g., Bierman, 1977; Kailath, 1980; Chin, 1983]. It is important to notice that proper mechanizations of the appropriate least-squares algorithms are definitely required for optimal efficiency and reliability of the results.

These considerations are especially critical in the context of firmware implementation of a formulation using Very Large Scale Integrated (VLSI) technology [Mead and Conway, 1980]. Such circuitry would correspond to complex arrays of subsystems to streamline the computations to their practical limits. For a wide spectrum of data processing applications, including on-line photogrammetry, these technological breakthroughs will undoubtedly revolutionize the optimization of data processing computations in the near future.

5. CONCLUDING REMARKS

The optimization of least-squares computations in on-line photogrammetry involves different aspects which require serious considerations for optimal results and maximum efficiency.

The adopted mathematical model for least-squares estimation purposes is clearly the most important decision that directly affects all the computations and hence the results. On a servo-driven photogrammetric system, the dynamics of the instrument should be considered jointly with the measurement process. This is particularly important when the on-line computations are carried out for the quality control of the observations.

The least-squares computations can be formulated in terms of covariance matrices, information matrices, orthogonal transformations or otherwise, without sacrificing any of the mathematical rigour of least-squares estimation. The appropriate selection of a formulation is therefore imperative to ensure computational efficiency, minimal storage requirements and optimal numerical accuracy.

In on-line photogrammetry as in other real-time application contexts, the mechanization of the selected algorithm in terms of software or firmware is also critical in view of the usual computer and time limitations. The optimization of the computations at this level of implementation requires much practical experience with the specific photogrammetric and computer systems. Further investigations are definitely required in these areas to reap the maximum benefits of on-line photogrammetric computations.

6. REFERENCES

AGEE, W.S. and R.H. TURNER (1972): Triangular Decomposition of a Positive Definite Matrix Plus a Symmetric Dyad With Application to Kalman Filtering.
Technical Report No. 38, White Sands Missile Range.

BIERMAN, G.J. (1977): Factorization Methods for Discrete Sequential Estimation.
Academic Press.

BLAIS, J.A.R. (1982): Recursive Least-Squares Estimation Using Givens Transformations.
Proceedings of the International Society of Photogrammetry and Remote Sensing, Commission III, Working Group 2, Symposium in Helsinki.

BLAIS, J.A.R. (1983): Linear Least-Squares Computations Using Givens Transformations. The Canadian Surveyor, Vol. 37, No. 4, pp. 225-233.

BLAIS, J.A.R. (1984): Least-Squares Filtering and Smoothing Using Givens Transformations. Under Preparation.

CHIN, L. (1983): Advances in Computational Efficiencies of Linear Filtering. Control and Dynamic Systems, pp. 125-192. Academic Press.

GELB, A. (ed.) (1974): Applied Optimal Estimation. The M.I.T. Press.

GENTLEMAN, W.M. (1973): Least-Squares Computations by Givens Transformations Without Square Roots. Journal of the Institute of Mathematical Applications, Vol. 12, pp. 329-336.

GEORGE, A. and M.T. Heath (1980): Solution of Sparse Linear Least-Squares Problems Using Givens Rotations. Linear Algebra and Its Applications, Vol. 34, pp. 69-83.

KAILATH, T. (1981): Lectures on Wiener and Kalman Filtering. CISM Courses and Lectures No. 140, Springer-Verlag.

KAMINSKI, P.G., A.E. BRYSON and S.F. SCHMIDT (1971): Discrete Square Root Filtering: A Survey of Current Techniques. IEEE Trans. Automatic Control AC-16, No. 6, pp. 727-735.

LAWSON, C.L. and R.J. HANSON (1974): Solving Least Squares Problems. Prentice-Hall.

MEAD, C. and L. CONWAY (1980): Introduction to VLSI Systems. Addison-Wesley Publishing Co.

TESKEY, T. (1984): Least-Squares Survey Computations Using Givens Transformations. M.Sc. Thesis, Publication No. 20005, Division of Surveying Engineering, The University of Calgary.