Block Cholesky Algorithms

Dr. Peter A. Steeves, P. Eng.
Geodetic Software Systems

A set of normal equations may be expressed in matrix notation as $N \cdot \hat{X} \;=\; B$.

If the normal equation matrix N is positive definite and symmetric, it is possible to find a triangular matrix [T] such that, $T^t \cdot T \;=\; N$, see [Ashkenazi, 1968]. The solution vector [X] and inverse [N^{-1}] can be found as follows:

$$T^t \cdot T \cdot \hat{X} \;=\; B$$

$$T \cdot \hat{X} \;=\; T^{t^{-1}} \cdot B$$

$$\hat{X} \;=\; T^{-1} \cdot T^{t^{-1}} \cdot B$$

$$N^{-1} \;=\; T^{-1} \cdot T^{t^{-1}}$$

Although these expressions involve the inverse of a triangular matrix, this is not required as it is possible to divide a matrix by a triangular matrix without computing the inverse.

Let $D \;=\; T \cdot \hat{X}$

then, $T^t \cdot D \;=\; B$

$$d_1 \;=\; b_1 / T^t_{1,1}$$

$$d_i \;=\; [\, b_i \;-\; \sum_{k=1}^{i-1} T^t_{i,k} \cdot d_k \,] \,/\, T^t_{i,i} \qquad i = 2, 3, 4, \;\ldots\ldots,\; n$$

$$T \cdot \hat{X} \;=\; D$$

$$x_n \;=\; d_n / T_{n,n}$$

$$x_i \;=\; [\, d_i \;-\; \sum_{k=i+1}^{n} T_{k,i} \cdot x_k \,] \,/\, T_{i,i} \qquad i = n\text{-}1, n\text{-}2, \;\ldots\ldots,\; 1$$

see [Steeves, 1974]. It has been demonstrated, [Knight and Steeves, 1974], that it is possible to find the inverse,

$$[N^{-1} = T^{-1} \cdot T^{t^{-1}}],$$

without computing the inverse of T. A substitution process begins with the last [*n*,*n*] element of the triangular matrix and proceeds row by row through the triangular matrix transforming it to N^{-1}. If B represents a lower square of the inverse inverse already available, then that part of the row directly above it and to the right of the diagonal element may be computed from,

$$b = \alpha \cdot t \cdot B$$

and the diagonal element from,

$$\beta = \alpha \cdot [\alpha - b \cdot t^t]$$

where, α is the inverse of the diagonal element and t is the respective row of T.

These algorithms were used to develop matrix partitioning schemes, where the algorithms are used to manipulate blocks of data rather than the unit approach as above. These matrix partitioning schemes will now be derived and then it will be shown how they may be used to solve sparse sets of normal equations by using control vectors.

Choleski Decomposition by Blocks:

Assume the normal equation matrix [N] and the triangular matrix [T], to be found, are partitioned into square blocks. The algorithm for the solution is:

$$T_{i,j} = N_{i,j} - \sum_{k=1}^{i-1} T^t_{k,i} \cdot T_{k,j}$$

$$T_{i,j} = T^{-1}_{i,i} \cdot T_{i,j} \qquad i \neq j$$

$$T_{i,i} = \sqrt{T_{i,i}} \qquad i = j$$

The solution involves computing the square root of a diagonal block and multiplying the inverse of a triangular block by a square block. Finding the square root of a block can be accomplished by using these same equations where each block becomes a single number. Multiplying the inverse of a triangular block by a block can be accomplished by the back substitution process.

Choleski Solution for Solution Vector by Blocks:

The algorithm is the same as that of the unit approach with the exception that each unit number now becomes a square block of numbers. Again the inverse of a triangular block multiplied by a block is accomplished by back substitution.

Choleski Solution for Inverse by Blocks:

The algorithm for the Choleski block inverse is as follows:

$$N^{-1}_{n,n} \;=\; T^{-1}_{n,n} \cdot T^{t^{-1}}_{n,n}$$

$i = n\text{-}1, n\text{-}2, \; . \; . \; . \; . \; . \; , \; 1 \qquad \{$

$$S_{i,j} \;=\; -T^{-1}_{i,i} \cdot \sum_{k=i+1}^{n} T_{i,k} \cdot S_{k,j} \qquad j = i+1, i+2, \; . \; . \; . \; . \; . \; , \; n$$

$$S_{i,i} \;=\; N^{-1}_{i,i} \;-\; T^{-1}_{i,i} \cdot \sum_{k=i+1}^{n} T_{i,k} \cdot S^{t}_{i,k} \quad \}$$

where,

$$N^{-1}_{i,i} \;=\; T^{-1}_{i,i} \cdot T^{t^{-1}}_{i,i} \;.$$

To compute $N^{-1}_{i,i}$ the inverse of the triangular block $[T_{i,i}]$ need not be found; the above algorithm may be used, where each block becomes a unit number.

Optimizing Sparseness by using Control Vectors:

A normal equation matrix might be very sparse, that is the number of non-zero elements may be only a small fraction of the total. If this is the case, then it may be possible to arrange the non-zero elements into some pattern symmetric about the principal axis of the matrix. Generally, in geomatic computations, this is the case and one may take advantage of it computationally wise. Methods of ordering normal equations into regular patterns may be found in [Ashkenazi, 1968] and [Snay, 1976].

Figure 1 denotes the normal equation matrix arising from the least squares adjustment of a horizontal control survey consisting of 257 stations. The non-zero elements of the matrix were partitioned into 46 [40 by 40] square blocks, the remainder of the matrix consisted of zeros. Each of the 46 blocks resided on a ramdon access data file and was called to the central processor when required.

The vector over the block matrix in Figure 1 is called a control vector and it denotes the number of zero block in each column of the normal equation matrix from the top down to the first non-zero block. The control vector controls the solution process so that no mathematical operations will be carried out within a zero block. The zero blocks are not even stored in the random access data file.

00	*00*	*00*	*01*	*02*	*03*	*04*	*05*	*06*	*07*	*07*	*09*	*01*
01	*02*	*04*										
	03	*05*	*07*									*35*
		06	*08*	*10*								*36*
			09	*11*	*13*							*37*
				12	*14*	*16*						*38*
					15	*17*	*19*					*39*
						18	*20*	*22*				*40*
							21	*23*	*25*	*28*		*41*
								24	*26*	*29*		*42*
									27	*30*	*32*	*43*
										31	*33*	*44*
											34	*45*
												46

Figure 1. Sequencing of Normal Equation
Partitions for Random Access Data File

The algorithm for the Choleski decomposition was given as,

$$T_{i,j} = N_{i,j} - \sum_{k=1}^{i-1} T^{t}_{k,i} \cdot T_{k,j}$$

$$T_{i,j} = T^{-1}_{i,i} \cdot T_{i,j} \qquad i \neq j$$

$$T_{i,i} = \sqrt{T_{i,i}} \qquad i = j$$

Assume Figure 1 represents the normal equation matrix and the first two block of the fifth column are to transformed to blocks of the triangular matrix.

$$T_{1,5} \quad = \quad N_{1,5} \quad - \quad \sum_{k=1}^{0} T^t_{k,i} \cdot T_{k,j}$$

$$T_{1,5} \quad = \quad 0 \quad - \quad 0 \quad = \quad 0$$

$$T_{2,5} \quad = \quad N_{2,5} \quad - \quad \sum_{k=1}^{1} T^t_{k,i} \cdot T_{k,j}$$

$$T_{2,5} \quad = \quad 0 \quad - \quad T_{1,2} \cdot 0$$

$$T_{2,5} \quad = \quad 0 \quad - \quad 0 \quad = \quad 0$$

It can also be seen that the reduced triangular matrix [T] will have the same form as [N] and, therefore, the same control vector may be used. Since, this is true we may replace

$$\sum_{k=1}^{i-1} \quad \text{with} \quad \sum_{k=l}^{i-1} \quad \text{where,}$$

l = Maximum (control vector [i], control vector [j]) and if control vector [j] $\geq$ i - 1 skip the algorithm and begin with the next block.

The control vector controls the number of computations, does not allow computation with zero blocks to be done, by setting the range of the inner product loop.

The same procedure is used for the Choleski solution vector and inverse.

It was shown, Knight and Steeves [1974], that it is possible to compute that part of the inverse corresponding to the original normal equation pattern without stepping outside of the variable band pattern.

Even though a great percentage of the zero entries have been eliminated, most of the non-zero blocks will have a number of zero in them. The computations involved in products, additions, square roots and back substitutions between these blocks can again be lessened if each block has its own control vector. For example, suppose we have two blocks U and V and we want to form the product:

$$Q \quad = \quad U^t \cdot V$$

the algorithm is as follows:
$i = 1, 2, 3, \; . \; . \; . \; . \; . \; , n$
$j = 1, 2, 3, \; . \; . \; . \; . \; . \; , n$

$$Q_{i,j} = 0$$

$$Q_{i,j} = Q_{i,j} + \sum_{k=1}^{n} U^{t}_{k,i} \cdot V_{k,j}$$

If the block *U* and *V* have control vectors, *IU* and *IV*, the summation could be changed to $\sum_{k=m}^{n}$ where, $m = maximum\ (IV[i], IV[j]]$.

References and Bibliography

Ashkenazi, V. Solution and error analysis of large geodetic networks, *Survey Review* No. 147, 1968.

Knight, Wm. and Steeves, P. Partial Solution of the Variance-Covariance Matrix for Geodetic Networks, The Canadian Surveyor, Vol. 28, No. 5, December 1974.

Steeves, P. Least Squares Adjustment of Horizontal Control on a Mapping Plane, Master of Science Thesis, University of New Brunswick, Fredericton, New Brunswick, 1974.

Snay, R. A. Reducing the profile of sparse symmetric matrices. Bull. Geod. 50 (4), pp. 341-352, 1976.