A block matrix generalization of Gauss-Jordan elimination using Haynsworth's quotient formula for Schur complements

Brad Pierce D. Stott Parker
pierce@cs.ucla.edu stott@cs.ucla.edu
Computer Science Department
University of California
Los Angeles, CA 90095-1596

Abstract

A block matrix algorithm for linear systems solution is derived using Haynsworth's quotient formula for Schur complements. The algorithm is essentially a generalization of Gauss-Jordan elimination.

Definition 1 *Let* $M = \begin{matrix} \\ k \\ n-k \end{matrix} \begin{matrix} \begin{matrix} k & m-k \end{matrix} \\ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \end{matrix}$ *with* A *nonsingular. Then* $(\ M\ /\ A\ ) = D - CA^{-1}B$ *is the* **Schur complement** *of* A *with respect to* M.

It may be easily verified that

$$\begin{pmatrix} I & 0 \\ CA^{-1} & (\ M\ /\ A\ ) \end{pmatrix} \begin{pmatrix} A & B \\ 0 & I \end{pmatrix} \;=\; M \;=\; \begin{pmatrix} A & 0 \\ C & I \end{pmatrix} \begin{pmatrix} I & A^{-1}B \\ 0 & (\ M\ /\ A\ ) \end{pmatrix}.$$

Hence, as was first remarked [1] by Schur [7, p. 217],

$$\det M \;=\; \det A \;\det(\ M\ /\ A\ ).$$

In particular, if M is nonsingular, then so is $(\ M\ /\ A\ )$.

If M has a $p \times q$ block decomposition (A_{ij}), then $(\ M\ /\ A_{11}\ )$ has the $(p-1) \times (q-1)$ block decomposition $\left((\begin{pmatrix} A_{11} & A_{1,j+1} \\ A_{i+1,1} & A_{i+1,j+1} \end{pmatrix} / A_{11}\ ) \right)$. For example, if $M = \begin{pmatrix} A & B & C \\ D & E & F \\ G & H & J \end{pmatrix}$, then

$$\begin{aligned} (\ M\ /\ A\ ) \;&=\; \begin{pmatrix} (\begin{pmatrix} A & B \\ D & E \end{pmatrix} / A\ ) & (\begin{pmatrix} A & C \\ D & F \end{pmatrix} / A\ ) \\ (\begin{pmatrix} A & B \\ G & H \end{pmatrix} / A\ ) & (\begin{pmatrix} A & C \\ G & J \end{pmatrix} / A\ ) \end{pmatrix} \\ &=\; \begin{pmatrix} E - DA^{-1}B & F - DA^{-1}C \\ H - GA^{-1}B & J - GA^{-1}C \end{pmatrix}. \end{aligned}$$

[1] This is according to Cottle [1, p. 190].

Theorem 1 *(Haynsworth) Let* $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} E & F \\ G & H \end{pmatrix} & B \\ C & D \end{pmatrix}$ *with* A *and* E *nonsingular. Then* $(\ A\ /\ E\ )$ *is nonsingular and* $(\ M\ /\ A\ ) = (\ (\ M\ /\ E\ )\ /\ (\ A\ /\ E\ )\ )$.

Proof
See [4]. This identity was first proved[2] by Haynsworth in [3]. □

Definition 2 $M[\,1..p\,|\,1..q\,]$ *is the upper-left* $p \times q$ *submatrix of* M. $M\langle k\rangle = M[\,1..k\,|\,1..k\,]$.

It is an immediate consequence of Haynsworth's quotient formula (together with Schur's determinantal identity) that:

Corollary 1 *Let* $k_1, k_2, \ldots, k_q$ *such that, for all* $1 \le i \le q$, $M\left\langle \sum_{1\le j\le i} k_j \right\rangle$ *is nonsingular. Define* $M^{(0)} = M$ *and, for* $1 \le i \le q$,

$$M^{(i)} \;=\; (\ M^{(i-1)}\ /\ M^{(i-1)}\langle k_i\rangle\ ).$$

Then $(\ M\ /\ M\left\langle \sum_{1\le j\le q} k_j \right\rangle\ ) = M^{(q)}$.

Let us now consider the special case of calculating $A^{-1}B$ by applying Haynsworth's method of successive Schur complements to $(\ M\ /\ A\ )$, where

$$M \;=\; \begin{array}{c} \\ n \\ n \end{array} \begin{array}{c} \begin{array}{cc} n & p \end{array} \\ \begin{pmatrix} A & B \\ -I & 0 \end{pmatrix} \end{array}.$$

Partitioning M as

$$\begin{array}{c} \\ k_1 \\ k_2 \\ \vdots \\ k_q \\ k_1 \\ k_2 \\ \vdots \\ k_q \end{array} \begin{array}{c} \begin{array}{cccc} k_1 & k_2 & \cdots & k_q \end{array} \\ \begin{pmatrix} A_{11}^{(0)} & A_{12}^{(0)} & \cdots & A_{1q}^{(0)} \\ A_{21}^{(0)} & A_{22}^{(0)} & \cdots & A_{2q}^{(0)} \\ \vdots & \vdots & \ddots & \vdots \\ A_{q1}^{(0)} & A_{q2}^{(0)} & \cdots & A_{qq}^{(0)} \\ -I & 0 & \cdots & 0 \\ 0 & -I & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots \\ 0 & 0 & \cdots & -I \end{pmatrix} \end{array} \begin{array}{c} p \\ \begin{pmatrix} B_1^{(0)} \\ B_2^{(0)} \\ \vdots \\ B_q^{(0)} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \end{array}$$

and calculating $(\ M\ /\ A_{11}^{(0)}\ )$ yields

$$\begin{array}{c} \\ k_2 \\ k_3 \\ \vdots \\ k_q \\ k_1 \\ k_2 \\ k_3 \\ \vdots \\ k_q \end{array} \begin{array}{c} \begin{array}{cccc} k_2 & k_3 & \cdots & k_q \end{array} \\ \begin{pmatrix} A_{22}^{(1)} & A_{23}^{(1)} & \cdots & A_{2q}^{(1)} \\ A_{32}^{(1)} & A_{33}^{(1)} & \cdots & A_{3q}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ A_{q2}^{(1)} & A_{q3}^{(1)} & \cdots & A_{qq}^{(1)} \\ A_{12}^{(1)} & A_{13}^{(1)} & \cdots & A_{1q}^{(1)} \\ -I & 0 & \cdots & 0 \\ 0 & -I & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots \\ 0 & 0 & \cdots & -I \end{pmatrix} \end{array} \begin{array}{c} p \\ \begin{pmatrix} B_2^{(1)} \\ B_3^{(1)} \\ \vdots \\ B_q^{(1)} \\ B_1^{(1)} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \end{array}$$

[2] This is according to Ostrowski [4, p. 389].

Continuing, ((M / $A_{11}^{(0)}$) / $A_{22}^{(1)}$) yields

$$\begin{array}{c} \\ k_3 \\ k_4 \\ \vdots \\ k_q \\ k_1 \\ k_2 \\ k_3 \\ k_4 \\ \vdots \\ k_q \end{array} \begin{array}{c} \begin{array}{ccccc} k_3 & k_4 & \cdots & k_q & p \end{array} \\ \begin{pmatrix} A_{33}^{(2)} & A_{23}^{(2)} & \cdots & A_{2q}^{(2)} & B_3^{(2)} \\ A_{32}^{(2)} & A_{33}^{(2)} & \cdots & A_{3q}^{(2)} & B_4^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A_{q2}^{(2)} & A_{q3}^{(2)} & \cdots & A_{qq}^{(2)} & B_q^{(2)} \\ A_{12}^{(2)} & A_{13}^{(2)} & \cdots & A_{1q}^{(2)} & B_1^{(2)} \\ A_{22}^{(2)} & A_{23}^{(2)} & \cdots & A_{2q}^{(2)} & B_2^{(2)} \\ -I & 0 & \cdots & 0 & 0 \\ 0 & -I & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -I & 0 \end{pmatrix} \end{array}$$

and so on, until eventually we terminate with ((((M / $A_{11}^{(0)}$) / $A_{22}^{(1)}$) / ...) / $A_{qq}^{(q-1)}$), yielding

$$\begin{array}{c} \\ k_1 \\ k_2 \\ \vdots \\ k_q \end{array} \begin{array}{c} \begin{array}{c} p \end{array} \\ \begin{pmatrix} B_1^{(q)} \\ B_2^{(q)} \\ \vdots \\ B_q^{(q)} \end{pmatrix} \end{array} \quad = \quad A^{-1}B$$

Note that only the first n rows of each iterate need be stored explicitly, hence

$$A^{-1}B \quad = \quad (\Xi_{k_q} \circ \ldots \circ \Xi_{k_2} \circ \Xi_{k_1})\,(\,A \quad B\,),$$

where Ξ is defined as follows:

Definition 3 *If* $K = \begin{array}{c} \\ k \\ n-k \end{array} \begin{array}{c} \begin{array}{cc} k & (n-k)+p \end{array} \\ \begin{pmatrix} E & F \\ G & H \end{pmatrix} \end{array}$ *with* E *nonsingular,* $1 \le k < n$, *and* $0 < p$, *then*

$$\Xi_k(K) \quad = \quad (\,\begin{pmatrix} E & F \\ G & H \\ -I_k & 0 \end{pmatrix} / \,E\,) \quad = \quad \begin{array}{c} \\ n-k \\ k \end{array} \begin{array}{c} \begin{array}{c} (n-k)+p \end{array} \\ \begin{pmatrix} H - GE^{-1}F \\ E^{-1}F \end{pmatrix} \end{array}.$$

If $K = \begin{array}{c} \\ n \end{array} \begin{array}{c} \begin{array}{cc} n & p \end{array} \\ \begin{pmatrix} E & F \end{pmatrix} \end{array}$ *with* E *nonsingular and* $0 < p$, *then* $\Xi_n(K) = E^{-1}F$.

Note that $\Xi_k(\begin{pmatrix} E & F \\ G & H \end{pmatrix}) = \begin{pmatrix} 0 & E \\ I_{n-k} & G \end{pmatrix}^{-1} \begin{pmatrix} F \\ H \end{pmatrix}$.

Theorem 2 *If* $(\Xi_{k_2} \circ \Xi_{k_1})(M)$ *is well-defined, then* $(\Xi_{k_2} \circ \Xi_{k_1})(M) = \Xi_{k_1+k_2}(M)$.

Proof

Let $M = \begin{array}{c} \\ k_1 \\ k_2 \end{array} \begin{array}{c} \begin{array}{ccc} k_1 & k_2 & \end{array} \\ \begin{pmatrix} E & F & B_1 \\ G & H & B_2 \\ C_1 & C_2 & D \end{pmatrix} \end{array}$ with E nonsingular and $A = \begin{pmatrix} E & F \\ G & H \end{pmatrix}$ nonsingular.

$$\Xi_{k_1+k_2}(M) = \begin{pmatrix} (\ M \ / \ A\ ) \\ A^{-1}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \end{pmatrix}$$

$$\begin{aligned}
(\Xi_{k_2} \circ \Xi_{k_1}) \begin{pmatrix} E & F & B_1 \\ G & H & B_2 \\ C_1 & C_2 & D \end{pmatrix} &= \Xi_{k_2} \begin{pmatrix} (\ A\ /\ E\ ) & B_2 - GE^{-1}B_1 \\ C_2 - C_1E^{-1}F & D - C_1E^{-1}B_1 \\ E^{-1}F & E^{-1}B_1 \end{pmatrix} \\
&= \begin{pmatrix} (\begin{pmatrix} (\ A\ /\ E\ ) & B_2 - GE^{-1}B_1 \\ C_2 - C_1E^{-1}F & D - C_1E^{-1}B_1 \end{pmatrix} / (\ A\ /\ E\ )) \\ (\begin{pmatrix} (\ A\ /\ E\ ) & B_2 - GE^{-1}B_1 \\ E^{-1}F & E^{-1}B_1 \end{pmatrix} / (\ A\ /\ E\ )) \\ (\ A\ /\ E\ )^{-1}(B_2 - GE^{-1}B_1) \end{pmatrix} \\
&= \begin{pmatrix} ((\ M\ /\ E\ ) / (\ A\ /\ E\ )) \\ E^{-1}B_1 - E^{-1}F(\ A\ /\ E\ )^{-1}(B_2 - GE^{-1}B_1) \\ (\ A\ /\ E\ )^{-1}(B_2 - GE^{-1}B_1) \end{pmatrix} \\
&= \begin{pmatrix} (\ M\ /\ A\ ) \\ A^{-1}\begin{pmatrix} E & F \\ G & H \end{pmatrix} Q \end{pmatrix},
\end{aligned}$$

where $Q = \begin{pmatrix} E^{-1}B_1 - E^{-1}F(\ A\ /\ E\ )^{-1}(B_2 - GE^{-1}B_1) \\ (\ A\ /\ E\ )^{-1}(B_2 - GE^{-1}B_1) \end{pmatrix}$.

$$\begin{aligned}
\begin{pmatrix} E & F \\ G & H \end{pmatrix} Q &= \begin{pmatrix} B_1 \\ GE^{-1}B_1 - GE^{-1}F(\ A\ /\ E\ )^{-1}(B_2 - GE^{-1}B_1) \\ \quad + H(\ A\ /\ E\ )^{-1}(B_2 - GE^{-1}B_1) \end{pmatrix} \\
&= \begin{pmatrix} B_1 \\ GE^{-1}B_1 + (H - GE^{-1}F)(\ A\ /\ E\ )^{-1}(B_2 - GE^{-1}B_1) \end{pmatrix} \\
&= \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}
\end{aligned}$$

□

The algorithm can be applied most flexibly when the nonsingular matrix A is also *Gauss nondegenerate*:

Definition 4 *A $p \times q$ matrix M of rank n is* **Gauss nondegenerate** *if, for all $1 \le k \le n$, $M\langle k \rangle$ is nonsingular.*

It is well-known [2, p.96] that if A is a nonsingular matrix, then Gaussian elimination (without augmentation by pivoting) is well-defined iff A is Gauss nondegenerate. We have shown elsewhere [5, 6] how input randomization can be used to efficiently achieve Gauss nondegeneracy (with probability 1) given any nonsingular matrix.

Theorem 3 *Let M be a Gauss nondegenerate $n \times n$ matrix. For all $1 \le k < n$, $(M \,/\, M\langle k \rangle)$ is Gauss nondegenerate.*

Proof

Let $1 \le k < n$, $M = \begin{array}{c} \\ k \\ n-k \end{array} \begin{array}{c} \begin{array}{cc} k & n-k \end{array} \\ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \end{array}$, and $1 \le j \le (n-k)$.

$$\begin{aligned} (\, M \,/\, A \,)\langle j \rangle &= D\langle j \rangle - (C A^{-1} B)\langle j \rangle \\ &= D\langle j \rangle - (C[\,1..j \,|\, 1..k\,]) A^{-1} (B[\,1..k \,|\, 1..j\,]) \\ &= (\, \begin{pmatrix} A & B[\,1..k \,|\, 1..j\,] \\ C[\,1..j \,|\, 1..k\,] & D\langle j \rangle \end{pmatrix} / \, A \,) \\ &= (\, M\langle k+j \rangle \,/\, A \,) \end{aligned}$$

Because M is Gauss nondegenerate, $M\langle k+j \rangle$ is nonsingular, hence so, too, is $(\, M\langle k+j \rangle \,/\, A \,)$. □

If A is a Gauss nondegenerate $n \times n$ matrix and B is an $n \times p$ matrix with $p > 1$, then the **Gauss-Jordan** elimination algorithm can be defined naturally in terms of Ξ. Initially $(k = 0)$,

$$\begin{array}{c} \\ n \end{array} \begin{array}{c} \begin{array}{cc} n & p \end{array} \\ \begin{pmatrix} A^{(0)} & B^{(0)} \end{pmatrix} \end{array} = (\, A \quad B \,).$$

For the first $n-1$ iterations $(k = 1, \ldots, n-1)$,

$$\begin{array}{c} \\ n \end{array} \begin{array}{c} \begin{array}{cc} n-k & p \end{array} \\ \begin{pmatrix} A^{(k)} & B^{(k)} \end{pmatrix} \end{array} = \begin{pmatrix} K_3^{(k)} & K_4^{(k)} \\ K_1^{(k)} & K_2^{(k)} \end{pmatrix},$$

where

$$\begin{aligned} \begin{array}{c} n-k \\ k \end{array} \begin{array}{c} \begin{array}{cc} n-k & p \end{array} \\ \begin{pmatrix} K_1^{(k)} & K_2^{(k)} \\ K_3^{(k)} & K_4^{(k)} \end{pmatrix} \end{array} &= \Xi_k \, (\, A \quad B \,) \\ &= (\, \overbrace{\Xi_1 \circ \cdots \circ \Xi_1}^{k \text{ times}} \,)(\, A \quad B \,). \end{aligned}$$

The final iteration $(k = n)$, yields $A^{-1}B$,

$$\begin{aligned} B^{(n)} &= \Xi_n \, (\, A \quad B \,) \\ &= (\, \overbrace{\Xi_1 \circ \cdots \circ \Xi_1}^{n \text{ times}} \,)(\, A \quad B \,). \end{aligned}$$

References

[1] R.W. Cottle. Manifestations of the Schur complement. *Linear Algebra Appl.*, 8:189–211, 1974.

[2] G. H. Golub and C. F. Van Loan. *Matrix Computations: second edition.* Johns Hopkins University Press, Baltimore, Maryland, 1989.

[3] E.V. Haynsworth. On the Schur complement. Basel Mathematical Notes 20, June 1968.

[4] A. Ostrowski. A new proof of Haynsworth's quotient formula for Schur complements. *Linear Algebra Appl.*, 4:389–392, 1971.

[5] D.S. Parker. Random butterfly transformations with applications in computational linear algebra. Technical Report CSD-950023, UCLA Computer Science Department, 1995.

[6] D.S. Parker and B. Pierce. The randomizing FFT: an alternative to partial pivoting in Gaussian elimination. Technical Report CSD-950037, UCLA Computer Science Department, 1995.

[7] I. Schur. Potenzreihen im Innern des Einheitskreises. *J. Reine Angew. Math.*, 147:205–232, 1917.