P

Partitioned Matrices and the Schur Complement

TABLE OF CONTENTS

Partitioned matrices often appear in the exposition of Finite Element Methods. This Appendix collects some basic material on the subject. Emphasis is placed on results involving the *Schur complement*. This is the name given in the linear algebra literature to matrix objects obtained through the condensation (partial elimination) process discussed in Chapter 10.

§P.1. Partitioned Matrix

Suppose that the square matrix $\mathbf{M}$ dimensioned $(n+m) \times (n+m)$, is partitioned into four submatrix blocks as

$$\underset{(n+m)\times(n+m)}{\mathbf{M}} = \begin{bmatrix} \underset{n\times n}{\mathbf{A}} & \underset{n\times m}{\mathbf{B}} \\ \underset{m\times n}{\mathbf{C}} & \underset{m\times m}{\mathbf{D}} \end{bmatrix} \tag{P.1}$$

The dimensions of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are as shown in the block display. $\mathbf{A}$ and $\mathbf{D}$ are square matrices, but $\mathbf{B}$ and $\mathbf{C}$ are not square unless $n = m$. Entries are generally assumed to be complex.

§P.2. Schur Complements

If $\mathbf{A}$ is nonsingular, the Schur complement of $\mathbf{M}$ with respect to $\mathbf{A}$ is defined as

$$\mathbf{M}/\mathbf{A} \stackrel{\text{def}}{=} \mathbf{D} - \mathbf{C}\,\mathbf{A}^{-1}\,\mathbf{B}. \tag{P.2}$$

If $\mathbf{D}$ is nonsingular, the Schur complement of $\mathbf{M}$ with respect to $\mathbf{D}$ is defined as

$$\mathbf{M}/\mathbf{D} \stackrel{\text{def}}{=} \mathbf{A} - \mathbf{B}\,\mathbf{D}^{-1}\,\mathbf{C}. \tag{P.3}$$

Matrices (P.2) and (P.3) are also called the *Schur complement of* $\mathbf{A}$ *in* $\mathbf{M}$ and the *Schur complement of* $\mathbf{D}$ *in* $\mathbf{M}$, respectively. These equivalent statements are sometimes preferable. If both diagonal blocks are singular, see §P.10. See **Notes and Bibliography** for other notations in common use.

§P.3. Block Diagonalization

The following *block diagonalization forms*, due to Aitken, clearly display the Schur complement role. If $\mathbf{A}$ is nonsingular,

$$\begin{gathered} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}\,\mathbf{A}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I} & -\mathbf{A}^{-1}\,\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}/\mathbf{A} \end{bmatrix}, \\ \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C}\,\mathbf{A}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}/\mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A}^{-1}\,\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \end{gathered} \tag{P.4}$$

whereas if $\mathbf{D}$ is nonsingular

$$\begin{gathered} \begin{bmatrix} \mathbf{I} & -\mathbf{B}\,\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{D}^{-1}\,\mathbf{C} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{M}/\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}, \\ \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{B}\,\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{M}/\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{D}^{-1}\,\mathbf{C} & \mathbf{I} \end{bmatrix}. \end{gathered} \tag{P.5}$$

All of these can be directly verified by matrix multiplication.

§P.4. Determinant Formulas

Taking determinants on both sides of the second of (P.4) yields

$$\det(\mathbf{M}) = \begin{vmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C}\,\mathbf{A}^{-1} & \mathbf{I} \end{vmatrix} \begin{vmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}/\mathbf{A} \end{vmatrix} \begin{vmatrix} \mathbf{I} & \mathbf{A}^{-1}\,\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{vmatrix} = \begin{vmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}/\mathbf{A} \end{vmatrix} = \det(\mathbf{A}).\det(\mathbf{M}/\mathbf{A}). \tag{P.6}$$

since the determinant of unit triangular matrices is one. Similarly, from the second of (P.4) one gets

$$\det(\mathbf{M}) = \det(\mathbf{D}).\det(\mathbf{M}/\mathbf{D}). \tag{P.7}$$

It follows that the determinant is multiplicative in the Schur complement. Extensions to the case in which $\mathbf{A}$ or $\mathbf{D}$ are singular, and its historical connection to Schur's paper [658] are discussed in [817].

§P.5. Partitioned Inversion

Suppose that the upper diagonal block $\mathbf{A}$ in (P.1) is nonsingular. Then the inverse of $\mathbf{M}$ in partitioned form may be expressed in several ways that involve (P.2):

$$\begin{aligned}
\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} &= \begin{bmatrix} \mathbf{I} & -\mathbf{A}^{-1}\,\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{M}/\mathbf{A})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}\,\mathbf{A}^{-1} & \mathbf{I} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{I} & -\mathbf{A}^{-1}\,\mathbf{B}\,(\mathbf{M}/\mathbf{A})^{-1} \\ \mathbf{0} & (\mathbf{M}/\mathbf{A})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ -\mathbf{C}\,\mathbf{A}^{-1} & \mathbf{I} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{A}^{-1}\,\mathbf{B} \\ \mathbf{I} \end{bmatrix} (\mathbf{M}/\mathbf{A})^{-1} \,[\, -\mathbf{C}\,\mathbf{A}^{-1} \quad \mathbf{I}\,] \\
&= \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\,\mathbf{B}\,(\mathbf{M}/\mathbf{A})^{-1}\,\mathbf{C}\,\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\,(\mathbf{M}/\mathbf{A})^{-1} \\ -(\mathbf{M}/\mathbf{A})^{-1}\,\mathbf{C}\,\mathbf{A}^{-1} & (\mathbf{M}/\mathbf{A})^{-1} \end{bmatrix}
\end{aligned} \tag{P.8}$$

Suppose next that the lower diagonal block $\mathbf{D}$ in (P.1) is nonsingular. Then the form corresponding to the last one above reads

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} (\mathbf{M}/\mathbf{D})^{-1} & -(\mathbf{M}/\mathbf{D})^{-1}\,\mathbf{B}\,\mathbf{A}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\,(\mathbf{M}/\mathbf{D})^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\,\mathbf{C}\,(\mathbf{M}/\mathbf{D})^{-1}\,\mathbf{B}\,\mathbf{D}^{-1} \end{bmatrix} \tag{P.9}$$

If both $\mathbf{A}$ and $\mathbf{D}$ are nonsingular, equating the top left hand block of (P.9) to that of the last form of (P.8) gives Duncan's inversion formula

$$(\mathbf{M}/\mathbf{D})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\,\mathbf{B}\,(\mathbf{M}/\mathbf{A})^{-1}\,\mathbf{C}\,\mathbf{A}^{-1}. \tag{P.10}$$

or explicitly in terms of the original blocks,

$$(\mathbf{A} - \mathbf{B}\,\mathbf{D}^{-1}\,\mathbf{C})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\,\mathbf{B}\,(\mathbf{D} - \mathbf{C}\,\mathbf{A}^{-1}\,\mathbf{B})^{-1}\,\mathbf{C}\,\mathbf{A}^{-1}. \tag{P.11}$$

When $n = m$ so that $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are all square, and all of them are nonsingular, one obtains the elegant formula due to Aitken:

$$\mathbf{M}^{-1} = \begin{bmatrix} (\mathbf{M}/\mathbf{D})^{-1} & (\mathbf{M}/\mathbf{B})^{-1} \\ (\mathbf{M}/\mathbf{C})^{-1} & (\mathbf{M}/\mathbf{A})^{-1} \end{bmatrix} \tag{P.12}$$

Replacing $-\mathbf{D}$ by the inverse of a matrix, say, $\mathbf{T}^{-1}$, leads to the Woodbury formula presented in Appendix D. If $\mathbf{B}$ and $\mathbf{C}$ are column and row vectors, respectively, vectors and $\mathbf{D}$ a scalar, one obtains the Sherman-Morrison inverse-update formula also presented in that Appendix.

§P.6. Solution of Partitioned Linear System

An important use of Schur complements is the partitioned solution of linear systems. In fact this is the most important application in FEM, in connection with superelement analysis. Suppose that we have the linear system $\mathbf{M}\,\mathbf{z} = \mathbf{r}$, in which the given coefficient matrix $\mathbf{M}$ is partitioned as in (P.1). The RHS $\mathbf{r} = [\,\mathbf{p}\ \ \mathbf{q}\,]^T$ is given, while $\mathbf{z} = [\,\mathbf{x}\ \ \mathbf{y}\,]^T$ is unknown. The system is equivalent to the matrix equation pair

$$\begin{aligned} \mathbf{A}\,\mathbf{x} + \mathbf{B}\,\mathbf{y} &= \mathbf{p}, \\ \mathbf{C}\,\mathbf{x} + \mathbf{D}\,\mathbf{y} &= \mathbf{q}. \end{aligned} \tag{P.13}$$

If $\mathbf{D}$ is nonsingular, eliminating $\mathbf{y}$ and solving for $\mathbf{x}$ yields

$$\mathbf{x} = (\mathbf{A} - \mathbf{B}\,\mathbf{D}^{-1}\,\mathbf{C})^{-1}\,(\mathbf{p} - \mathbf{B}\,\mathbf{D}^{-1}\mathbf{q}) = (\mathbf{M}/\mathbf{D})^{-1}\,(\mathbf{p} - \mathbf{B}\,\mathbf{D}^{-1}\mathbf{q}). \tag{P.14}$$

Similarly if $\mathbf{A}$ is nonsingular, eliminating $\mathbf{x}$ and solving for $\mathbf{y}$ yields

$$\mathbf{y} = (\mathbf{D} - \mathbf{C}\,\mathbf{A}^{-1}\,\mathbf{B})^{-1}\,(\mathbf{q} - \mathbf{C}\,\mathbf{A}^{-1}\mathbf{p}) = (\mathbf{M}/\mathbf{A})^{-1}\,(\mathbf{q} - \mathbf{C}\,\mathbf{A}^{-1}\mathbf{p}). \tag{P.15}$$

Readers familiar with the process of static condensation in FEM will notice that (P.14) is precisely the process of elimination of internal degrees of freedom of a superelement, as discussed in Chapter 10. In this case $\mathbf{M}$ becomes the superelement stiffness matrix $\mathbf{K}$, which is symmetric. In the notation of that Chapter, the appropriate block mapping is $\mathbf{A} \to \mathbf{K}_{bb}$, $\mathbf{B} \to \mathbf{K}_{bi}$, $\mathbf{C} \to \mathbf{K}_{ib} = \mathbf{K}_{bi}^T$, $\mathbf{D} \to \mathbf{K}_{ii}$, $\mathbf{p} \to \mathbf{f}_b$, $\mathbf{q} \to \mathbf{f}_i$, $\mathbf{x} \to \mathbf{u}_b$, and $\mathbf{y} \to \mathbf{u}_i$. The Schur complement $\mathbf{M}/\mathbf{D} \to \mathbf{K}/\mathbf{K}_{ii}$ is the condensed stiffness matrix $\tilde{\mathbf{K}}_{bb} = \mathbf{K}_{bb} - \mathbf{K}_{bi}\,\mathbf{K}_{ii}^{-1}\,\mathbf{K}_{ib}$, while the RHS of (P.14) is the condensed force vector $\tilde{\mathbf{f}}_b = \mathbf{f}_b - \mathbf{K}_{bi}\,\mathbf{K}_{ii}^{-1}\,\mathbf{f}_i$.

§P.7. Rank Additivity

From (P.4) or (P.5) we immediately obtain the *rank additivity* formulas

$$\mathrm{rank}(\mathbf{M}) = \mathrm{rank}(\mathbf{A}) + \mathrm{rank}(\mathbf{M}/\mathbf{A}) = \mathrm{rank}(\mathbf{D}) + \mathrm{rank}(\mathbf{M}/\mathbf{D}), \tag{P.16}$$

or, explicitly in term of the original blocks

$$\mathrm{rank}\left(\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}\right) = \mathrm{rank}(\mathbf{A}) + \mathrm{rank}(\mathbf{D} - \mathbf{B}\,\mathbf{A}^{-1}\mathbf{C}) = \mathrm{rank}(\mathbf{D}) + \mathrm{rank}(\mathbf{A} - \mathbf{C}\,\mathbf{D}^{-1}\mathbf{B}), \tag{P.17}$$

assuming that the indicated inverses exist. Consequently the rank is additive in the Schur complement, and so it the inertia of a Hermitian matrix, as considered next.

§P.8. Inertia Additivity

Consider the (generally complex) partitioned Hermitian matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^* & \mathbf{A}_{22} \end{bmatrix}. \tag{P.18}$$

Here $\mathbf{A}_{11}$ is assumed nonsingular, and $\mathbf{A}_{12}^*$ denotes the conjugate transpose of $\mathbf{A}_{12}$.

The *inertia* of $\mathbf{A}$ is a list of three nonnegative integers

$$\mathrm{In}(\mathbf{A}) = \{n_+, n_-, n_0\}, \tag{P.19}$$

in which $n_+ = n_+(\mathbf{A})$, $n_- = n_-(\mathbf{A})$, and $n_0 = n_0(\mathbf{A})$ give the number of positive, negative, and zero eigenvalues of $\mathbf{A}$, respectively. (Recall that all eigenvalues of a Hermitian matrix are real.) Since $\mathbf{A}$ is Hermitian, its rank is $n_+(\mathbf{A}) + n_-(\mathbf{A})$, and its nullity $n_0(\mathbf{A})$. (Note that n_0 is the dimension of the kernel of $\mathbf{A}$ and also its corank.) The *inertia additivity formula* in terms of the Schur complement of $\mathbf{A}_{11}$ in $\mathbf{A}$, is

$$\mathrm{In}(\mathbf{A}) = \mathrm{In}(\mathbf{A}_{11}) + \mathrm{In}(\mathbf{A}/\mathbf{A}_{11}). \tag{P.20}$$

It follows that if $\mathbf{A}_{11}$ is positive definite, and $\mathbf{A}/\mathbf{A}_{11}$ is positive (nonnegative) definite, then $\mathbf{A}$ is positive (nonnegative) definite. Also $\mathbf{A}$ and $\mathbf{A}/\mathbf{A}_{11}$ have the same nullity.

§P.9. The Quotient Property

Consider the partitioned matrix

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \vdots & \mathbf{B}_1 \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \vdots & \mathbf{B}_2 \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{C}_1 & \mathbf{C}_2 & \vdots & \mathbf{D} \end{bmatrix}, \tag{P.21}$$

in which both $\mathbf{A}$ and $\mathbf{A}_{11}$ are nonsingular. Matrix $\mathbf{M}$ need not be square. The *quotient property* of nested Schur complements is

$$\mathbf{M}/\mathbf{A} = (\mathbf{M}/\mathbf{A}_{11})/(\mathbf{A}/\mathbf{A}_{11}). \tag{P.22}$$

On the RHS of (P.22) we see that the "denominator" matrix $\mathbf{A}_{11}$ "cancels out" as if the expressions were scalar fractions. If $\mathbf{M}$ is square and nonsingular, then taking determinants and aplying the determinant formula we obtain

$$\det(\mathbf{M}/\mathbf{A}) = \det(\mathbf{M}/\mathbf{A}_{11})/\det(\mathbf{A}/\mathbf{A}_{11}). \tag{P.23}$$

§P.10. Generalized Schur Complements

If both diagonal blocks $\mathbf{A}$ and $\mathbf{D}$ of $\mathbf{M}$ in (P.1) are singular, the notion of Schur complement may be generalized by replacing the conventional inverse with the Moore-Penrose generalized inverse. Further developments of this fairly advanced topic may be found in Chapters 1 and 6 of [817].

Notes and Bibliography

Most of the following historical remarks are extracted from the **Historical Introduction** chapter of [817]. This edited book gives a comprehensive and up-to-date coverage of the Schur's complement and its use as a basic tool in linear algebra.

The name "Schur complement" for the matrices defined in (P.2) and (P.3) was introduced by Haynsworth in 1968 [345,346]. The name was motivated by the seminal "determinant lemma" by Schur published in 1917 in his studies of stability polynomials [658]. Earlier implicit manifestations of this idea starting with Laplace in 1812 are described in the **Historical Introduction** chapter of [817].

Although (P.6) and (P.7) are now commonly referred to as *Schur determinant formulas* in the literature, he actually proved only a special case, in which $n = m$ so all block matrices in (P.1) are square; in addition $\mathbf{B}$ and $\mathbf{C}$ were required to commute. The effect of these assumptions is that the formula can be extended to singular diagonal blocks. See §0.3 of [817] for details.

The Schur complement appears naturally in the inversion, linear-equation solving, rank and inertia determinations that involve block partitioned matrices. Among the most seminal formulas are:

- The last of the partitioned inversion formulas (P.8) was first published in 1937 by Banachievicz [44]. It appeared one year later (independently derived) in the classical monograph by Frazer, Duncan and Collar [282], cited in Appendix H as one of the early sources for Matrix Structural Analysis.
- The Duncan inversion formula (P.10) appeared in 1944 [192]. It is a follow up to the inversion methods of [282].
- The Aitken block diagonalization formulas (P.4) and (P.5) appear in another classic book from the same period [7].
- The inertia additivity formula (P.20) as well as the quotient property (P.22) are due to Haynsworth and coworkers [345,346]; see also [156,347,533,534]
- From the Duncan inversion formula follows directly the Guttman rank additivity formula (P.16) first published in 1946 [322], as well as the already cited inertia additivity formula for Hermitian matrices (P.20), which appeared much later [345,346].

The notation $\mathbf{M}/\mathbf{A}$ for $\mathbf{D} - \mathbf{C}\,\mathbf{A}^{-1}\,\mathbf{B}$ is now commonly used in linear algebra. Others are $(\mathbf{M}/\mathbf{A})$, $\mathbf{M}|\mathbf{A}$, $(\mathbf{M}|\mathbf{A})$, and $\mathbf{M}_A$. The latter has visualization problems if matrix blocks are subscripted; for example $\mathbf{M}_{A_{22}}$ is clumsier to read than $\mathbf{M}/\mathbf{A}_{22}$ or $\mathbf{M}|\mathbf{A}_{22}$. A drawback of the slash (and vertical bar) notation is that enclosing parentheses are necessary if one takes a transpose or inverse; e.g., compare $(\mathbf{M}/\mathbf{A}_{22})^{-1}$ versus the more compact $\mathbf{M}^{-1}_{A_{22}}$.