Section 3.4. More on Partitioned Square Matrices: The Schur Complement

Note. In this section, we associate a quantity with a partitioned matrix and express the inverse (when it exists) and the determinant of a matrix in terms of this quantity.

Note. Recall that a full rank partitioning of a square $n \times n$ matrix A of rank r is $A = \begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$ where W is an $r \times r$ full rank matrix, X is $r \times (n-r)$, Y is $(n-r) \times r$, and Z is $(n-r) \times (n-r)$. So $[W\,X]$ is of full row rank r and the rows of $[W\,X]$ span the row space of A. Also, $\begin{bmatrix} W \\ Y \end{bmatrix}$ is of full column rank r and the columns of $\begin{bmatrix} W \\ Y \end{bmatrix}$ span the column space of A. So the rows of $[Y\,Z]$ can be written as linear combinations of the rows of $[W\,X]$ and so there is some $(n-r) \times r$ matrix T such that $[Y\,Z] = T[W\,X]$. Similarly, there is $r \times (n-r)$ matrix S such that $\begin{bmatrix} X \\ Z \end{bmatrix} = \begin{bmatrix} W \\ Y \end{bmatrix} S$. So we have $Y = TW$, $Z = TX$, $X = WS$, and $Z = YS$, so that $Z = TX = TWS$. Since W is of full rank, then W^{-1} exists so that $T = YW^{-1}$, $S = W^{-1}X$, and $Z = YS = YW^{-1}X$ (or, equivalently, $Z = TX = YW^{-1}X$). So a full rank partitioning can be written in terms of W, X, and Y only as

$$A = \begin{bmatrix} W & X \\ Y & YW^{-1}X \end{bmatrix}. \qquad (*)$$

Definition. If A is a square matrix partitioned as $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ where A_{11} is nonsingular, then $Z = A_{22} - A_{21}A_{11}^{-1}A_{12}$ is the *Schur complement* of A_{11} in A.

Note. If A_{11} is of full rank and $\text{rank}(A_{11}) = \text{rank}(A)$, then from $(*)$ we see that $A_{22} = A_{21}A_{11}^{-1}A_{12}$ and so in this case $Z = 0$.

Note. As described in the note after Theorem 3.3.6, for any $n \times m$ matrix A of rank $r > 0$, there is a $n \times n$ permutation matrix E_{π_1} and a $m \times m$ permutation matrix E_{π_2} such that $E_{\pi_1}AE_{\pi_2}$ can be partitioned as $E_{\pi_1}AE_{\pi_2} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ where B_{11} is a $r \times r$ full rank matrix. So from $(*)$ we have $E_{\pi_1}AE_{\pi_2} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{21}B_{11}^{-1}B_{12} \end{bmatrix}$. We can then factor as:

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} [I \;\; B_{11}^{-1}B_{12}]$$

$$= \begin{bmatrix} I \\ B_{21}B_{11}^{-1} \end{bmatrix} B_{11}[I \;\; B_{11}^{-1}B_{12}] = \begin{bmatrix} I \\ B_{21}B_{11}^{-1} \end{bmatrix} [B_{11} \;\; B_{12}].$$

With $P = E_{\pi_1}^{-1}$ and $Q = E_{\pi_2}^{-1}$ we have

$$A = P \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} [I \;\; B_{11}^{-1}B_{12}]Q = \begin{bmatrix} PB_{11} \\ PB_{21} \end{bmatrix} [Q \;\; B_{11}^{-1}B_{12}Q] \qquad (**)$$

and

$$A = P \begin{bmatrix} I \\ B_{21}B_{11}^{-1} \end{bmatrix} [B_{11} \;\; B_{12}]Q = \begin{bmatrix} P \\ PB_{21}B_{11}^{-1} \end{bmatrix} [B_{11}Q \;\; B_{12}Q]. \qquad (***)$$

Now P and Q are permutation matrices and so are of full rank and B_{11} is of full rank, so $\begin{bmatrix} PB_{11} \\ PB_{21} \end{bmatrix}$ and $\begin{bmatrix} P \\ PB_{21}B_{11}^{-1} \end{bmatrix}$ are of full column rank and $[Q \;\; B_{11}^{-1}B_{12}Q]$ and $[B_{11}Q \;\; B_{12}Q]$ are of full row rank. So $(**)$ and $(***)$ give two factorizations of A in the form $A = LR$ where L is a $n \times r$ full column rank matrix and R is a $r \times m$ full row rank.

Definition. If $n \times m$ matrix A of rank r can be written as $A = LR$ where L is a $n \times r$ full column rank matrix and R is a $r \times m$ full row rank matrix, then $A = LR$ is a *full rank factorization* of A.

Theorem 3.4.1. If A is a square nonsingular matrix and $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ where both A_{11} and A_{22} are nonsingular then in terms of the Schur complement of A_{11} in A, $Z = A_{22} - A_{21}A_{11}^{-1}A_{12}$, we have that the inverse of A is

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}Z^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}Z^{-1} \\ -Z^{-1}A_{21}A_{11}^{-1} & Z^{-1} \end{bmatrix}.$$

Note. The proof of Theorem 3.4.1 is to be given in Exercise 3.13.

Theorem 3.4.2. If A is a square matrix such that $A = \begin{bmatrix} X^T \\ y^T \end{bmatrix} [X \; y]$ where X is of full column rank, then the Schur complement of X^TX in A is

$$y^Ty - y^TX(X^TX)^{-1}X^Ty.$$

Theorem 3.4.3. If A is a square matrix partitioned as $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ where A_{11} is square and nonsingular then

$$\det(A) = \det(A_{11})\det(A_{22} - A_{21}A_{11}^{-1}A_{12}) = \det(A_{11})\det(Z)$$

where $Z = A_{22} - A_{21}A_{11}^{-1}A_{12}$ is the Schur complement of A_{11} in A.

Revised: 7/1/2018