

8.3 Schur Complements

Section 1 inverted block triangular matrices, and Section 2 cleared blocks without changing the rank. A general block matrix $\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ has no zero block to help, but when one diagonal block is invertible, a single step of block elimination creates one. What is left in the opposite corner is a new matrix, the Schur complement, and it carries everything that the invertible block does not: the determinant, the rank, and the inverse of the whole matrix can all be read off from the block and its complement. This section introduces it, derives these formulas, uses the two complements of one matrix to prove the Sherman–Morrison–Woodbury formula, and ends with the determinant of a block matrix whose lower blocks commute, over every field.

8.3.1 Eliminating a block

Suppose we must solve a linear system whose unknowns come in two groups, $\mathbf{x} \in F^k$ and $\mathbf{y} \in F^l$:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix}, \quad \text{that is,} \quad \begin{aligned} \mathbf{Ax} + \mathbf{By} &= \mathbf{f}, \\ \mathbf{Cx} + \mathbf{Dy} &= \mathbf{g}, \end{aligned}$$

with $\mathbf{A} \in M_k(F)$ invertible. Do what we would do for two equations in two numbers: solve the first equation for $\mathbf{x}$, and substitute. The first equation gives $\mathbf{x} = \mathbf{A}^{-1}(\mathbf{f} - \mathbf{By})$. Substituting into the second,

$$\mathbf{CA}^{-1}\mathbf{f} - \mathbf{CA}^{-1}\mathbf{By} + \mathbf{Dy} = \mathbf{g}, \quad \text{so} \quad (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})\mathbf{y} = \mathbf{g} - \mathbf{CA}^{-1}\mathbf{f}.$$

The unknown $\mathbf{x}$ has disappeared, and $\mathbf{y}$ solves a smaller system whose matrix is $\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$. Once $\mathbf{y}$ is known, $\mathbf{x} = \mathbf{A}^{-1}(\mathbf{f} - \mathbf{By})$. So the big system is solvable for every right-hand side exactly when the small matrix $\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$ allows it, and this matrix will keep appearing.

For $k = l = 1$ we have met it before: $d - cb/a = (ad - bc)/a$ is the determinant divided by the pivot a , which is what Gaussian elimination leaves in the corner after clearing the entry c .

The Schur complement of an invertible block is what remains in the opposite corner after that block has been used to eliminate its row and column.

Definition 8.3.1. Schur Complement

Let $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \in M_{(k+p) \times (k+l)}(F)$ be a block matrix with $\mathbf{A} \in M_k(F)$, $\mathbf{B} \in M_{k \times l}(F)$, $\mathbf{C} \in M_{p \times k}(F)$, and $\mathbf{D} \in M_{p \times l}(F)$, where $k, p, l \geq 1$.

(a) If $\mathbf{A}$ is invertible, the **Schur complement of $\mathbf{A}$ in $\mathbf{M}$** is

$$\mathbf{M}/\mathbf{A} := \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \in M_{p \times l}(F).$$

(b) Symmetrically, if $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ with $\mathbf{D} \in M_l(F)$ invertible, $\mathbf{A} \in M_{m \times k}(F)$, $\mathbf{B} \in M_{m \times l}(F)$, and $\mathbf{C} \in M_{l \times k}(F)$, the **Schur complement of $\mathbf{D}$ in $\mathbf{M}$** is $\mathbf{M}/\mathbf{D} := \mathbf{A} - \mathbf{BD}^{-1}\mathbf{C} \in M_{m \times k}(F)$.

In words, clause by clause. The block whose complement we take must be **square and invertible**, since its inverse appears in the formula; the other blocks may be rectangular. The

complement has the size of the **opposite corner**: $\mathbf{M}/\mathbf{A}$ has the shape of $\mathbf{D}$, and $\mathbf{M}/\mathbf{D}$ has the shape of $\mathbf{A}$. In $\mathbf{CA}^{-1}\mathbf{B}$ the factors come in the only order the sizes allow: $\mathbf{C}$ is $p \times k$, $\mathbf{A}^{-1}$ is $k \times k$, $\mathbf{B}$ is $k \times l$. A way to remember it: go from the corner $\mathbf{D}$ to $\mathbf{A}$ along its row ($\mathbf{C}$), invert, and come back along its column ($\mathbf{B}$).

Well-definedness. $\mathbf{M}/\mathbf{A}$ depends on $\mathbf{M}$ and on the chosen partition, which says which corner is $\mathbf{A}$. The inverse $\mathbf{A}^{-1}$ is unique (Theorem 0.9.7), so there is no other choice to make.

Examples.

- **Scalars.** For $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $a \neq 0$, $\mathbf{M}/a = d - ca^{-1}b = (ad - bc)/a$. In particular $a \cdot (\mathbf{M}/a) = \det \mathbf{M}$.
- **An identity corner.** If $\mathbf{A} = \mathbf{I}_k$, then $\mathbf{M}/\mathbf{I}_k = \mathbf{D} - \mathbf{CB}$. No inversion is needed at all.
- **Block triangular.** If $\mathbf{C} = \mathbf{O}$, then $\mathbf{M}/\mathbf{A} = \mathbf{D}$: nothing needs to be eliminated.
- **A 3×3 example.** For $\mathbf{M} = \left(\begin{array}{cc|c} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 3 & 1 & 4 \end{array} \right)$ over $\mathbb{Q}$, $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ has $\mathbf{A}^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$,

and $\mathbf{CA}^{-1} = \begin{pmatrix} 3 & 1 \end{pmatrix} \mathbf{A}^{-1} = \begin{pmatrix} 2 & -1 \end{pmatrix}$, so $\mathbf{M}/\mathbf{A} = 4 - \begin{pmatrix} 2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = (2)$. This 1×1 complement is the last pivot of Gaussian elimination on $\mathbf{M}$: subtracting $\frac{1}{2}$ row 1 from row 2 and $\frac{3}{2}$ row 1 from row 3 gives the rows $(0, \frac{1}{2}, -\frac{1}{2})$ and $(0, -\frac{1}{2}, \frac{5}{2})$, and adding the first of these to the second gives $(0, 0, 2)$.

Nonexample by minimal change. For $\mathbf{M} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $\mathbf{M}/(1) = 0 - 1 \cdot 1 \cdot 1 = (-1)$. Change the corner entry 1 to 0: $\mathbf{M}' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. The blocks $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$, and the partition are unchanged, and $\mathbf{M}'$ is still invertible, but $\mathbf{M}'/\mathbf{O}_1$ is **not defined**, because the clause “ $\mathbf{A}$ is invertible” fails. Here $\mathbf{M}'/\mathbf{D}$ is not defined either, since $\mathbf{D} = \mathbf{O}_1$.

Why this definition. The formula is read off from the elimination in the hook, so the question is only about the name and the hypotheses. “Complement” refers to position: $\mathbf{M}/\mathbf{A}$ sits in the corner complementary to $\mathbf{A}$. The quotient notation $\mathbf{M}/\mathbf{A}$ is suggested by the first example, $\det(\mathbf{M}/a) = \det \mathbf{M} / \det a$, and the determinant formula below shows that this holds for blocks of every size. Invertibility of $\mathbf{A}$ cannot be dropped, since $\mathbf{A}^{-1}$ is in the formula; the price is that some invertible matrices, like $\mathbf{M}'$, have no Schur complement for a given partition.

Warning. $\mathbf{M}$ can be invertible while $\mathbf{A}$ is not. For $\mathbf{M}' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, neither corner is invertible, so neither Schur complement exists for the partition 1, 1. Before using a formula below, check that the block you invert is invertible. If it is not, try the other corner ($\mathbf{M}/\mathbf{D}$), a different partition, or first permute rows: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{M}' = \mathbf{I}_2$. The same issue appears in Gaussian elimination as a zero pivot.

Quick check. Let $\mathbf{M} = \left(\begin{array}{c|cc} 2 & 4 & 2 \\ \hline 1 & 3 & 5 \\ 3 & 1 & 0 \end{array} \right)$ with partition 1, 2. Compute $\mathbf{M}/(2)$.

Solution. Here $\mathbf{A} = (2)$, $\mathbf{B} = \begin{pmatrix} 4 & 2 \end{pmatrix}$, $\mathbf{C} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 3 & 5 \\ 1 & 0 \end{pmatrix}$. Then $\mathbf{CA}^{-1}\mathbf{B} = \frac{1}{2} \begin{pmatrix} 4 & 2 \\ 12 & 6 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix}$, so $\mathbf{M}/(2) = \begin{pmatrix} 1 & 4 \\ -5 & -3 \end{pmatrix}$. These are exactly the entries left in rows 2, 3 after clearing column 1 with the pivot 2.

8.3.2 The block LDU factorization

The elimination of the hook, written with matrices, factors $\mathbf{M}$ into a lower block triangular matrix, a block diagonal matrix, and an upper block triangular matrix. Everything else in this section follows from this factorization.

Theorem 8.3.1. Block LDU Factorization

Let $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ be partitioned as in Definition 8.3.1.

(a) If $\mathbf{A} \in M_k(F)$ is invertible, then

$$\mathbf{M} = \begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ \mathbf{CA}^{-1} & \mathbf{I}_p \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{M}/\mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{I}_k & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix}.$$

(b) If $\mathbf{D} \in M_l(F)$ is invertible, then

$$\mathbf{M} = \begin{pmatrix} \mathbf{I}_m & \mathbf{BD}^{-1} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix} \begin{pmatrix} \mathbf{M}/\mathbf{D} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ \mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_l \end{pmatrix}.$$

Idea. The left factor in (a) records the row operation “subtract $\mathbf{CA}^{-1}$ times block row 1 from block row 2”, which turns $\mathbf{C}$ into $\mathbf{O}$ and $\mathbf{D}$ into $\mathbf{M}/\mathbf{A}$. The right factor records the column operation “subtract block column 1 times $\mathbf{A}^{-1}\mathbf{B}$ from block column 2”, which turns $\mathbf{B}$ into $\mathbf{O}$. Undoing both operations gives the factorization. To prove it we need not retrace the elimination: multiplying out is enough.

Proof.

(a) Write $\mathbf{S} = \mathbf{M}/\mathbf{A} = \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$. The sizes are conformable throughout. By Theorem 8.1.1,

$$\begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{S} \end{pmatrix} \begin{pmatrix} \mathbf{I}_k & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{AA}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{S} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{S} \end{pmatrix},$$

and then

$$\begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ \mathbf{CA}^{-1} & \mathbf{I}_p \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{S} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{CA}^{-1}\mathbf{A} & \mathbf{CA}^{-1}\mathbf{B} + \mathbf{S} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix},$$

since $\mathbf{CA}^{-1}\mathbf{A} = \mathbf{C}$ and $\mathbf{CA}^{-1}\mathbf{B} + (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}) = \mathbf{D}$. By associativity (Theorem 0.9.4) the triple product equals $\mathbf{M}$.

(b) Write $\mathbf{S}' = \mathbf{M}/\mathbf{D} = \mathbf{A} - \mathbf{BD}^{-1}\mathbf{C}$. By Theorem 8.1.1,

$$\begin{pmatrix} \mathbf{S}' & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ \mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_l \end{pmatrix} = \begin{pmatrix} \mathbf{S}' & \mathbf{O} \\ \mathbf{C} & \mathbf{D} \end{pmatrix},$$

$$\begin{pmatrix} \mathbf{I}_m & \mathbf{BD}^{-1} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix} \begin{pmatrix} \mathbf{S}' & \mathbf{O} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{S}' + \mathbf{BD}^{-1}\mathbf{C} & \mathbf{BD}^{-1}\mathbf{D} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \mathbf{M}.$$

■

The outer factors are block triangular with identity blocks on the diagonal. They are invertible, with inverses obtained by changing the sign of the off-diagonal block (Theorem 8.1.5), and they have determinant 1. So, up to these harmless factors, $\mathbf{M}$ is the block diagonal matrix $\mathbf{A} \oplus (\mathbf{M}/\mathbf{A})$ when $\mathbf{A}$ and $\mathbf{D}$ are square. Determinant, rank, and invertibility cannot tell the difference.

8.3.3 Determinant and rank

The outer factors of Theorem 8.3.1 do not change the determinant or the rank, so both can be read off from the middle factor.

Theorem 8.3.2. Schur Determinant Formula

Let $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \in M_{k+l}(F)$ with $\mathbf{A} \in M_k(F)$, $\mathbf{D} \in M_l(F)$, and $k, l \geq 1$.

- (a) If $\mathbf{A}$ is invertible, then $\det \mathbf{M} = \det \mathbf{A} \cdot \det(\mathbf{M}/\mathbf{A})$.
- (b) If $\mathbf{D}$ is invertible, then $\det \mathbf{M} = \det \mathbf{D} \cdot \det(\mathbf{M}/\mathbf{D})$.

Proof.

- (a) By Theorem 8.3.1 (a), $\mathbf{M} = \mathbf{L}(\mathbf{A} \oplus (\mathbf{M}/\mathbf{A}))\mathbf{U}$ with $\mathbf{L} = \begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ \mathbf{CA}^{-1} & \mathbf{I}_l \end{pmatrix}$ and $\mathbf{U} = \begin{pmatrix} \mathbf{I}_k & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix}$. Here $p = l$ because $\mathbf{M}$ is square, so $\mathbf{M}/\mathbf{A} \in M_l(F)$. By Theorem 7.8.4, $\det \mathbf{L} = \det \mathbf{I}_k \det \mathbf{I}_l = 1$, $\det \mathbf{U} = 1$, and $\det(\mathbf{A} \oplus (\mathbf{M}/\mathbf{A})) = \det \mathbf{A} \det(\mathbf{M}/\mathbf{A})$. By Theorem 7.5.4,

$$\det \mathbf{M} = \det \mathbf{L} \cdot \det \mathbf{A} \det(\mathbf{M}/\mathbf{A}) \cdot \det \mathbf{U} = \det \mathbf{A} \det(\mathbf{M}/\mathbf{A}).$$

- (b) The same argument with Theorem 8.3.1 (b), where $\mathbf{M}/\mathbf{D} \in M_k(F)$ because $\mathbf{M}$ is square.

■

This is the correct replacement for the false rule $\det \mathbf{M} = \det(\mathbf{AD} - \mathbf{BC})$ of Section 1: the determinant of $\mathbf{M}$ is that of $\mathbf{A}$ times that of the complement. For $k = l = 1$ it is $a(d - ca^{-1}b) = ad - bc$. Chapter 7 promised this, and it relies on Theorem 7.8.4 in exactly the form proved there, with square diagonal blocks.

Theorem 8.3.3. Schur Rank Formula

Let $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ be partitioned as in Definition 8.3.1. If $\mathbf{A} \in M_k(F)$ is invertible, then

$$\text{rank } \mathbf{M} = k + \text{rank}(\mathbf{M}/\mathbf{A}).$$

Symmetrically, if $\mathbf{D} \in M_l(F)$ is invertible, then $\text{rank } \mathbf{M} = l + \text{rank}(\mathbf{M}/\mathbf{D})$.

Proof. By Theorem 8.3.1 (a), $\mathbf{M} = \mathbb{L} \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{M}/\mathbf{A} \end{pmatrix} \mathbf{U}$, where $\mathbb{L}$ and $\mathbf{U}$ are invertible by Theorem 8.2.2 (a). By Theorem 3.5.9, $\text{rank } \mathbf{M}$ equals the rank of the middle factor, which is $\text{rank } \mathbf{A} + \text{rank}(\mathbf{M}/\mathbf{A}) = k + \text{rank}(\mathbf{M}/\mathbf{A})$ by Theorem 8.2.1 (b) and Theorem 3.4.4. The second statement follows in the same way from Theorem 8.3.1 (b). ■

In words: an invertible block of size k contributes exactly k to the rank, and all the remaining rank sits in its complement. For example, in Example (Rank of $\mathbf{I} - \mathbf{AB}$ versus $\mathbf{I} - \mathbf{BA}$) the matrix $\begin{pmatrix} \mathbf{I}_m & \mathbf{A} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix}$ has the two complements $\mathbf{I}_n - \mathbf{BA}$ and $\mathbf{I}_m - \mathbf{AB}$, and the theorem gives both expressions for its rank at once.

8.3.4 Inverting a block matrix

Since $\mathbf{A} \oplus (\mathbf{M}/\mathbf{A})$ is invertible exactly when $\mathbf{M}/\mathbf{A}$ is, the factorization also inverts $\mathbf{M}$: invert the three factors and multiply them in the reverse order.

Theorem 8.3.4. Block Inverse Formula

Let $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \in M_{k+l}(F)$ with $\mathbf{A} \in M_k(F)$ and $\mathbf{D} \in M_l(F)$.

- (a) Suppose $\mathbf{A}$ is invertible, and let $\mathbf{S} = \mathbf{M}/\mathbf{A}$. Then $\mathbf{M}$ is invertible if and only if $\mathbf{S}$ is invertible, and in that case

$$\mathbf{M}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{S}^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{S}^{-1} \end{pmatrix}.$$

In particular the lower-right $l \times l$ block of $\mathbf{M}^{-1}$ is $(\mathbf{M}/\mathbf{A})^{-1}$.

- (b) Suppose $\mathbf{D}$ is invertible, and let $\mathbf{S}' = \mathbf{M}/\mathbf{D}$. Then $\mathbf{M}$ is invertible if and only if $\mathbf{S}'$ is

invertible, and in that case

$$\mathbf{M}^{-1} = \begin{pmatrix} \mathbf{S}'^{-1} & -\mathbf{S}'^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{S}'^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}\mathbf{S}'^{-1}\mathbf{B}\mathbf{D}^{-1} \end{pmatrix}.$$

In particular the upper-left $k \times k$ block of $\mathbf{M}^{-1}$ is $(\mathbf{M}/\mathbf{D})^{-1}$.

Proof.

- (a) By [Theorem 8.3.3](#), $\text{rank } \mathbf{M} = k + \text{rank } \mathbf{S}$. So $\text{rank } \mathbf{M} = k + l$ if and only if $\text{rank } \mathbf{S} = l$, and by [Theorem 3.4.4](#) this says that $\mathbf{M}$ is invertible if and only if $\mathbf{S}$ is.

Suppose $\mathbf{S}$ is invertible. In [Theorem 8.3.1](#) (a), $\mathbf{M} = \mathbf{L}(\mathbf{A} \oplus \mathbf{S})\mathbf{U}$, where $\mathbf{L}^{-1} = \begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_l \end{pmatrix}$ and $\mathbf{U}^{-1} = \begin{pmatrix} \mathbf{I}_k & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix}$ by [Theorem 8.1.5](#), and $(\mathbf{A} \oplus \mathbf{S})^{-1} = \mathbf{A}^{-1} \oplus \mathbf{S}^{-1}$ by [Theorem 8.1.4](#) (c). By [Theorem 0.9.7](#) (3), $\mathbf{M}^{-1} = \mathbf{U}^{-1}(\mathbf{A}^{-1} \oplus \mathbf{S}^{-1})\mathbf{L}^{-1}$. By [Theorem 8.1.1](#),

$$\begin{aligned} \mathbf{U}^{-1}(\mathbf{A}^{-1} \oplus \mathbf{S}^{-1}) &= \begin{pmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{S}^{-1} \\ \mathbf{O} & \mathbf{S}^{-1} \end{pmatrix}, \\ \begin{pmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{S}^{-1} \\ \mathbf{O} & \mathbf{S}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_l \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{S}^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{S}^{-1} \end{pmatrix}. \end{aligned}$$

- (b) The equivalence follows from the second statement of [Theorem 8.3.3](#) in the same way.

By [Theorem 8.3.1](#) (b), $\mathbf{M}^{-1} = \begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_l \end{pmatrix} (\mathbf{S}'^{-1} \oplus \mathbf{D}^{-1}) \begin{pmatrix} \mathbf{I}_k & -\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix}$, and

$$\begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_l \end{pmatrix} \begin{pmatrix} \mathbf{S}'^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{D}^{-1} \end{pmatrix} = \begin{pmatrix} \mathbf{S}'^{-1} & \mathbf{O} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{S}'^{-1} & \mathbf{D}^{-1} \end{pmatrix},$$

whose product with $\begin{pmatrix} \mathbf{I}_k & -\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix}$ is the displayed matrix by [Theorem 8.1.1](#).

■

Nobody should memorize these blocks. Remember instead that $\mathbf{M}^{-1}$ is the product of three easy inverses in reverse order, and that the corner opposite $\mathbf{A}$ in $\mathbf{M}^{-1}$ is the inverse of the complement of $\mathbf{A}$.

Example. A 4×4 Inverse through a 2×2 Schur Complement

Invert

$$\mathbf{M} = \left(\begin{array}{cc|cc} 1 & 1 & 0 & -1 \\ 1 & 2 & -1 & -1 \\ \hline -1 & -1 & 1 & 1 \\ 0 & 1 & 0 & -1 \end{array} \right) \in M_4(\mathbb{Q})$$

using [Theorem 8.3.4](#), and find $\det \mathbf{M}$.

Solution. The blocks are $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix}$, $\mathbf{C} = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$.

Step 1: invert $\mathbf{A}$. $\det \mathbf{A} = 1$, so $\mathbf{A}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ by Theorem 0.9.8.

Step 2: the two “half” products.

$$\mathbf{CA}^{-1} = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}, \quad \mathbf{A}^{-1}\mathbf{B} = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{CA}^{-1}\mathbf{B} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Step 3: the complement and its inverse. $\mathbf{S} = \mathbf{M}/\mathbf{A} = \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$. Its determinant is $-1 \neq 0$, so $\mathbf{M}$ is invertible by Theorem 8.3.4 (a), and $\mathbf{S}^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} = \mathbf{S}$ by Theorem 0.9.8. By Theorem 8.3.2, $\det \mathbf{M} = \det \mathbf{A} \det \mathbf{S} = -1$.

Step 4: the blocks of $\mathbf{M}^{-1}$. Using Step 2,

$$\begin{aligned} -\mathbf{A}^{-1}\mathbf{BS}^{-1} &= -\begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \\ -\mathbf{S}^{-1}\mathbf{CA}^{-1} &= -\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \end{aligned}$$

and

$$\begin{aligned} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{BS}^{-1}\mathbf{CA}^{-1} &= \mathbf{A}^{-1} + (\mathbf{A}^{-1}\mathbf{BS}^{-1})(\mathbf{CA}^{-1}) \\ &= \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Hence

$$\mathbf{M}^{-1} = \left(\begin{array}{cc|cc} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ \hline 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & -1 \end{array} \right).$$

Check one row: row 1 of $\mathbf{M}$ times column 1 of $\mathbf{M}^{-1}$ is $1 + 0 + 0 + 0 = 1$, and times column 3 is $0 + 1 + 0 - 1 = 0$. All the work was with 2×2 matrices, and only two of them had to be inverted.

Remark. When $\mathbf{M}$ is real symmetric, $\mathbf{M}/\mathbf{A}$ is again symmetric, and Chapter 13 uses Schur complements to test whether $\mathbf{M}$ is positive definite one block at a time. Chapter 24 reads Gaussian elimination itself as a sequence of Schur complements of 1×1 pivots, which is the point of view behind block algorithms in numerical linear algebra.

8.3.5 The Sherman–Morrison–Woodbury formula

Chapter 7 proved the Matrix Determinant Lemma, $\det(\mathbf{I}_n + \mathbf{u}\mathbf{v}^\top) = 1 + \mathbf{v}^\top \mathbf{u}$ (Theorem 7.8.5), by bordering $\mathbf{I}_n + \mathbf{u}\mathbf{v}^\top$ into a larger block matrix. There is a matching statement about inverses: if we already know $\mathbf{A}^{-1}$ and change $\mathbf{A}$ by a low-rank term $\mathbf{UCV}$, the new inverse should cost only a small inversion. The trick is to find one block matrix that has $\mathbf{A}$ in one corner and $\mathbf{A} + \mathbf{UCV}$ as a Schur complement, and then to compute its inverse twice.

Theorem 8.3.5. Sherman–Morrison–Woodbury Formula

Let $n, k \geq 1$, let $\mathbf{A} \in M_n(F)$ and $\mathbf{C} \in M_k(F)$ be invertible, and let $\mathbf{U} \in M_{n \times k}(F)$ and $\mathbf{V} \in M_{k \times n}(F)$. Put $\mathbf{W} = \mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U} \in M_k(F)$.

- (a) $\mathbf{A} + \mathbf{UCV}$ is invertible if and only if $\mathbf{W}$ is invertible.
- (b) In that case

$$\begin{aligned} (\mathbf{A} + \mathbf{UCV})^{-1} &= \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}\mathbf{W}^{-1}\mathbf{VA}^{-1} \\ &= \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U})^{-1}\mathbf{VA}^{-1}. \end{aligned}$$

- (c) $\det(\mathbf{A} + \mathbf{UCV}) = \det \mathbf{A} \cdot \det \mathbf{C} \cdot \det \mathbf{W}$.

Idea. Look for a block matrix whose two Schur complements are $\mathbf{A} + \mathbf{UCV}$ and something involving $\mathbf{W}$. The complement of the lower-right corner of $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{U} \\ \mathbf{V} & \mathbf{D} \end{pmatrix}$ is $\mathbf{A} - \mathbf{UD}^{-1}\mathbf{V}$, which is $\mathbf{A} + \mathbf{UCV}$ when $\mathbf{D}^{-1} = -\mathbf{C}$, that is, $\mathbf{D} = -\mathbf{C}^{-1}$. The complement of the upper-left corner is then $-\mathbf{C}^{-1} - \mathbf{VA}^{-1}\mathbf{U} = -\mathbf{W}$. Now Theorem 8.3.4 computes the upper-left block of $\mathbf{M}^{-1}$ in two ways: as $(\mathbf{M}/\mathbf{D})^{-1} = (\mathbf{A} + \mathbf{UCV})^{-1}$, and from the formula in terms of $\mathbf{A}$ and $\mathbf{M}/\mathbf{A}$. Setting them equal is the Woodbury formula. The determinant formula of the previous subsection, used twice, gives (c).

Proof. Let $\mathbf{D} = -\mathbf{C}^{-1}$, which is invertible with $\mathbf{D}^{-1} = -\mathbf{C}$ by Theorem 0.9.7 (2) and (5), and let $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{U} \\ \mathbf{V} & \mathbf{D} \end{pmatrix} \in M_{n+k}(F)$. By Definition 8.3.1,

$$\begin{aligned} \mathbf{M}/\mathbf{A} &= \mathbf{D} - \mathbf{VA}^{-1}\mathbf{U} = -\mathbf{C}^{-1} - \mathbf{VA}^{-1}\mathbf{U} = -\mathbf{W}, \\ \mathbf{M}/\mathbf{D} &= \mathbf{A} - \mathbf{UD}^{-1}\mathbf{V} = \mathbf{A} + \mathbf{UCV}. \end{aligned}$$

- (a) By Theorem 8.3.4 (a) and (b), $\mathbf{M}$ is invertible if and only if $-\mathbf{W}$ is, and if and only if $\mathbf{A} + \mathbf{UCV}$ is. Since $-\mathbf{W}$ is invertible exactly when $\mathbf{W}$ is (Theorem 0.9.7 (5)), this proves (a).
- (b) Suppose $\mathbf{W}$ is invertible, so $\mathbf{M}$ is invertible, and $(-\mathbf{W})^{-1} = -\mathbf{W}^{-1}$. The upper-left $n \times n$ block of $\mathbf{M}^{-1}$ is $(\mathbf{M}/\mathbf{D})^{-1} = (\mathbf{A} + \mathbf{UCV})^{-1}$ by Theorem 8.3.4 (b). By Theorem 8.3.4 (a), applied to $\mathbf{M}$, whose blocks are $\mathbf{A}, \mathbf{U}, \mathbf{V}, \mathbf{D}$, and with $\mathbf{S} = \mathbf{M}/\mathbf{A} = -\mathbf{W}$, the same block is

$$\mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{U}(-\mathbf{W})^{-1}\mathbf{VA}^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}\mathbf{W}^{-1}\mathbf{VA}^{-1}.$$

The two expressions are the same block of the same matrix, so they are equal.

- (c) By [Theorem 8.3.2](#) (a) and (b), $\det \mathbf{A} \det(-\mathbf{W}) = \det \mathbf{M} = \det \mathbf{D} \det(\mathbf{A} + \mathbf{UCV})$. Multiplying a $k \times k$ matrix by -1 multiplies each of its k rows by -1 , so by [Theorem 7.5.1](#) $\det(-\mathbf{W}) = (-1)^k \det \mathbf{W}$ and $\det \mathbf{D} = (-1)^k \det(\mathbf{C}^{-1}) = (-1)^k (\det \mathbf{C})^{-1}$, using [Corollary 7.5.6](#). Hence $\det \mathbf{A} (-1)^k \det \mathbf{W} = (-1)^k (\det \mathbf{C})^{-1} \det(\mathbf{A} + \mathbf{UCV})$. Multiplying both sides by $(-1)^k \det \mathbf{C}$ gives (c). ■

The formula is useful when k is much smaller than n and $\mathbf{A}^{-1}$ is already known or easy (diagonal, say): the only new inversion is of the $k \times k$ matrix $\mathbf{W}$. The most common case is a rank-one update, $k = 1$ and $\mathbf{C} = (1)$.

Corollary 8.3.6. Sherman–Morrison Formula

Let $\mathbf{A} \in M_n(F)$ be invertible and $\mathbf{u}, \mathbf{v} \in F^n$. Then $\mathbf{A} + \mathbf{uv}^\top$ is invertible if and only if $1 + \mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u} \neq 0$, and in that case

$$(\mathbf{A} + \mathbf{uv}^\top)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1} \mathbf{u} \mathbf{v}^\top \mathbf{A}^{-1}}{1 + \mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u}}.$$

Moreover $\det(\mathbf{A} + \mathbf{uv}^\top) = \det \mathbf{A} (1 + \mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u})$.

Proof. Apply [Theorem 8.3.5](#) with $k = 1$, $\mathbf{C} = (1)$, $\mathbf{U} = \mathbf{u}$, and $\mathbf{V} = \mathbf{v}^\top$. Then $\mathbf{W} = (1 + \mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u})$ is a 1×1 matrix, invertible exactly when its entry is nonzero, with inverse the reciprocal. Multiplying by a 1×1 matrix is multiplying by its entry, which gives the displayed formula; and $\det \mathbf{W}$ is the entry itself. ■

For example, over $\mathbb{Q}$ let $\mathbf{A} = \text{diag}(1, -1)$, so $\mathbf{A}^{-1} = \mathbf{A}$, and $\mathbf{u} = (1, 1)$. For $\mathbf{v} = (2, 1)$, $\mathbf{A}^{-1} \mathbf{u} = (1, -1)$, $\mathbf{v}^\top \mathbf{A}^{-1} = \begin{pmatrix} 2 & -1 \end{pmatrix}$, and $1 + \mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u} = 1 + (2 - 1) = 2 \neq 0$. So $\mathbf{A} + \mathbf{uv}^\top = \begin{pmatrix} 3 & 1 \\ 2 & 0 \end{pmatrix}$ has determinant $\det \mathbf{A} \cdot 2 = -2$, and

$$\begin{aligned} (\mathbf{A} + \mathbf{uv}^\top)^{-1} &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} 2 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \begin{pmatrix} 1 & -\frac{1}{2} \\ -1 & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{3}{2} \end{pmatrix}, \end{aligned}$$

which agrees with [Theorem 0.9.8](#). For $\mathbf{v} = (1, 2)$ instead, $1 + \mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u} = 1 + (1 - 2) = 0$, and indeed $\mathbf{A} + \mathbf{uv}^\top = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}$ is not invertible.

With $\mathbf{A} = \mathbf{I}_n$, the determinant statement is the Matrix Determinant Lemma of Chapter 7, now obtained as a special case of a formula about general low-rank updates, and it matches Exercise (C3: Rank-one updates of an invertible matrix).

8.3.6 Commuting blocks

Section 1 warned that $\det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \neq \det(\mathbf{AD} - \mathbf{BC})$ in general. With the Schur complement we can see what goes wrong and what extra hypothesis repairs it. Take all four blocks in $M_n(F)$ and suppose $\mathbf{D}$ is invertible. By Theorem 8.3.2 (b),

$$\begin{aligned} \det \mathbf{M} &= \det \mathbf{D} \det(\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C}) \\ &= \det((\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})\mathbf{D}) = \det(\mathbf{AD} - \mathbf{BD}^{-1}\mathbf{CD}), \end{aligned}$$

using Theorem 7.5.4. If $\mathbf{C}$ and $\mathbf{D}$ commute, then $\mathbf{D}^{-1}\mathbf{CD} = \mathbf{D}^{-1}\mathbf{DC} = \mathbf{C}$, and the right side is $\det(\mathbf{AD} - \mathbf{BC})$. So commuting is exactly what lets $\mathbf{D}^{-1}$ cancel. The real question is how to remove the hypothesis that $\mathbf{D}$ is invertible, which is not needed for the conclusion to make sense.

Over $\mathbb{R}$ one could perturb: $\mathbf{D} + t\mathbf{I}_n$ is invertible for all but finitely many t , both sides are polynomials in t , and polynomials that agree at infinitely many points are equal. Over a finite field such as $\mathbb{F}_2$ there may be no invertible perturbation at all. We use instead a perturbation by an **indeterminate**: replace $\mathbf{D}$ by $\mathbf{D} + x\mathbf{I}_n$, a matrix with entries in $F[x]$. Its determinant is a monic polynomial, hence not the zero polynomial, and that is all the cancellation needs. First we record that evaluating polynomial entries commutes with matrix operations.

For $\mathbf{N} = (n_{ij})$ with entries in $F[x]$ and $c \in F$, write $\mathbf{N}(c) = (n_{ij}(c))$ for the matrix over F obtained by evaluating every entry at c . A matrix over F is a matrix over $F[x]$ with constant entries, and then $\mathbf{N}(c) = \mathbf{N}$.

Lemma 8.3.7. Evaluating a Matrix of Polynomials

Let $c \in F$, and let $\mathbf{N}, \mathbf{N}'$ be matrices with entries in $F[x]$ of sizes for which the operations below are defined. Then $(\mathbf{N} + \mathbf{N}')(c) = \mathbf{N}(c) + \mathbf{N}'(c)$, $(\mathbf{NN}')(c) = \mathbf{N}(c)\mathbf{N}'(c)$, and, for square $\mathbf{N}$, $(\det \mathbf{N})(c) = \det(\mathbf{N}(c))$.

Proof. By Theorem 0.8.5, evaluation at c sends sums of polynomials to sums, products to products, and constants to themselves. The (i, j) -entry of $\mathbf{N} + \mathbf{N}'$ is $n_{ij} + n'_{ij}$ and that of $\mathbf{NN}'$ is $\sum_l n_{il}n'_{lj}$, so evaluating gives the corresponding entries of $\mathbf{N}(c) + \mathbf{N}'(c)$ and $\mathbf{N}(c)\mathbf{N}'(c)$. The determinant $\det \mathbf{N}$ is the Leibniz sum $\sum_{\sigma} \text{sgn}(\sigma) \prod_j n_{\sigma(j)j}$ of Definition 7.4.1, computed in $F[x]$, and evaluating it term by term gives the Leibniz sum for $\mathbf{N}(c)$. This is the argument of Lemma 7.10.1. ■

Theorem 8.3.8. Determinant with Commuting Blocks

Let $n \geq 1$ and $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D} \in M_n(F)$ over any field F . If $\mathbf{CD} = \mathbf{DC}$, then

$$\det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \det(\mathbf{AD} - \mathbf{BC}).$$

Idea. ① Work over the commutative ring $F[x]$, replacing $\mathbf{D}$ by $\mathbf{D}_x = \mathbf{D} + x\mathbf{I}_n$, which still commutes with $\mathbf{C}$. ② Multiply $\mathbf{M}_x = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D}_x \end{pmatrix}$ on the right by $\begin{pmatrix} \mathbf{D}_x & \mathbf{O} \\ -\mathbf{C} & \mathbf{I}_n \end{pmatrix}$: the lower-left block becomes $\mathbf{CD}_x - \mathbf{D}_x\mathbf{C} = \mathbf{O}$ and the upper-left block becomes $\mathbf{AD}_x - \mathbf{BC}$. This is the Schur computation above with $\mathbf{D}^{-1}$ cleared from the denominators. ③ Take determinants: $\det \mathbf{M}_x \det \mathbf{D}_x = \det(\mathbf{AD}_x - \mathbf{BC}) \det \mathbf{D}_x$. ④ Cancel $\det \mathbf{D}_x$, which is legal because it is a nonzero polynomial and $F[x]$ has no zero divisors. ⑤ Set $x = 0$. Each step uses only determinant facts that Chapter 7 proved over commutative rings.

Proof. Regard $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ as matrices over $F[x]$ with constant entries, and let $\mathbf{D}_x = \mathbf{D} + x\mathbf{I}_n$. The proofs of Theorem 0.9.4 use only the ring axioms, so they hold for matrices over $F[x]$; in particular

$$\mathbf{CD}_x = \mathbf{CD} + x\mathbf{C} = \mathbf{DC} + x\mathbf{C} = \mathbf{D}_x\mathbf{C},$$

using the hypothesis $\mathbf{CD} = \mathbf{DC}$. Let

$$\mathbf{M}_x = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D}_x \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} \mathbf{D}_x & \mathbf{O} \\ -\mathbf{C} & \mathbf{I}_n \end{pmatrix},$$

both in $M_{2n}(F[x])$. By Theorem 8.1.1, which holds over $F[x]$ by the remark following it,

$$\mathbf{M}_x\mathbf{R} = \begin{pmatrix} \mathbf{AD}_x - \mathbf{BC} & \mathbf{B} \\ \mathbf{CD}_x - \mathbf{D}_x\mathbf{C} & \mathbf{D}_x \end{pmatrix} = \begin{pmatrix} \mathbf{AD}_x - \mathbf{BC} & \mathbf{B} \\ \mathbf{O} & \mathbf{D}_x \end{pmatrix}.$$

By the remarks in Chapter 7 §5 and §8, Theorem 7.5.4 and Theorem 7.8.4 hold over the commutative ring $F[x]$. Hence $\det(\mathbf{M}_x\mathbf{R}) = \det \mathbf{M}_x \det \mathbf{R}$, with $\det \mathbf{R} = \det \mathbf{D}_x \det \mathbf{I}_n = \det \mathbf{D}_x$, and $\det(\mathbf{M}_x\mathbf{R}) = \det(\mathbf{AD}_x - \mathbf{BC}) \det \mathbf{D}_x$. So, in $F[x]$,

$$\det \mathbf{M}_x \cdot \det \mathbf{D}_x = \det(\mathbf{AD}_x - \mathbf{BC}) \cdot \det \mathbf{D}_x. \quad (*)$$

Now $\mathbf{D}_x = x\mathbf{I}_n - (-\mathbf{D})$, so $\det \mathbf{D}_x = p_{-\mathbf{D}}(x)$ is the characteristic polynomial of $-\mathbf{D} \in M_n(F)$ (Definition 7.10.1). By Theorem 7.10.2 it is monic of degree n , so $\det \mathbf{D}_x \neq 0$ in $F[x]$. By Corollary 0.8.4 (2), we may cancel it from (*):

$$\det \mathbf{M}_x = \det(\mathbf{AD}_x - \mathbf{BC}) \quad \text{in } F[x].$$

Finally evaluate both sides at 0. By Lemma 8.3.7, $(\det \mathbf{M}_x)(0) = \det(\mathbf{M}_x(0)) = \det \mathbf{M}$, since $\mathbf{D}_x(0) = \mathbf{D}$; and $(\mathbf{AD}_x - \mathbf{BC})(0) = \mathbf{AD} - \mathbf{BC}$, so $(\det(\mathbf{AD}_x - \mathbf{BC}))(0) = \det(\mathbf{AD} - \mathbf{BC})$. Hence $\det \mathbf{M} = \det(\mathbf{AD} - \mathbf{BC})$. ■

The cancellation in $F[x]$ is the heart of the argument, so it is worth being precise about why it is legal. Over F we cannot divide by $\det \mathbf{D}$ when it is 0. Over $F[x]$ we still cannot divide, since $\det \mathbf{D}_x$ usually has no inverse, but we do not need to: cancellation only needs “ $pq = pr$ and $p \neq 0$ imply $q = r$ ”, and this holds in $F[x]$ because a product of nonzero polynomials is nonzero (Corollary 0.8.4). The perturbation $x\mathbf{I}_n$ makes $\det \mathbf{D}_x$ nonzero **as a polynomial** even when every value $\det(\mathbf{D} + c\mathbf{I}_n)$ with $c \in F$ is zero. For example, over $\mathbb{F}_2$, let $\mathbf{D} = \mathbf{E}_{22} = \text{diag}(0, 1)$. Then $\det(\mathbf{D} + 0\mathbf{I}_2) = 0$ and $\det(\mathbf{D} + \mathbf{I}_2) = \det \text{diag}(1, 0) = 0$, because $1 + 1 = 0$ in $\mathbb{F}_2$; no scalar shift makes $\mathbf{D}$ invertible. But $\det(\mathbf{D} + x\mathbf{I}_2) = x(x + 1) = x^2 + x$ is not the zero polynomial.

For example, over $\mathbb{Q}$, let $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, and $\mathbf{C} = 2\mathbf{I}_2 + \mathbf{D} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, which commutes with $\mathbf{D}$. Here $\mathbf{D}$ is not invertible, so the Schur route is closed. The theorem gives $\det \mathbf{M} = \det(\mathbf{AD} - \mathbf{BC}) = \det \begin{pmatrix} -2 & 0 \\ -2 & -3 \end{pmatrix} = 6$.

Warning. Which blocks commute decides which product appears. The hypothesis $\mathbf{CD} = \mathbf{DC}$ gives $\det(\mathbf{AD} - \mathbf{BC})$. If instead $\mathbf{A}$ and $\mathbf{C}$ commute, the correct formula is $\det(\mathbf{AD} - \mathbf{CB})$ (Exercise (C3: When $\mathbf{A}$ and $\mathbf{C}$ commute)), with $\mathbf{C}$ and $\mathbf{B}$ in the other order. The two can differ: for the blocks $\mathbf{A} = \mathbf{I}_2$, $\mathbf{B} = \mathbf{E}_{12}$, $\mathbf{C} = \mathbf{E}_{21}$, $\mathbf{D} = \mathbf{E}_{11}$ of the warning in Section 1, $\mathbf{A}$ commutes with $\mathbf{C}$ but $\mathbf{C}$ does not commute with $\mathbf{D}$; indeed $\det \mathbf{M} = -1 = \det(\mathbf{E}_{11} - \mathbf{E}_{22}) = \det(\mathbf{AD} - \mathbf{CB})$, while $\det(\mathbf{AD} - \mathbf{BC}) = 0$.

8.3.7 Exercises

8.3.7.1 A. Check your understanding

Exercise. A1

- Define the Schur complement $\mathbf{M}/\mathbf{A}$, including the hypothesis on $\mathbf{A}$ and the size of $\mathbf{M}/\mathbf{A}$.
- State the block LDU factorization of $\mathbf{M}$ in terms of $\mathbf{A}$ and $\mathbf{M}/\mathbf{A}$.
- True or false: if $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ is invertible and $\mathbf{A}$ is square, then $\mathbf{A}$ is invertible. Justify your answer.
- If $\mathbf{A}$ is invertible and $\mathbf{M}$ is invertible, which block of $\mathbf{M}^{-1}$ is $(\mathbf{M}/\mathbf{A})^{-1}$?
- True or false: for all $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D} \in M_n(F)$ with $\mathbf{CD} = \mathbf{DC}$, $\det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \det(\mathbf{AD} - \mathbf{BC})$, even if $\mathbf{D}$ is not invertible and $F = \mathbb{F}_2$. Justify your answer.

Solution.

- For $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ with $\mathbf{A} \in M_k(F)$ invertible, $\mathbf{B} \in M_{k \times l}(F)$, $\mathbf{C} \in M_{p \times k}(F)$, $\mathbf{D} \in M_{p \times l}(F)$: $\mathbf{M}/\mathbf{A} = \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \in M_{p \times l}(F)$, the size of $\mathbf{D}$ (Definition 8.3.1).
- $\mathbf{M} = \begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ \mathbf{CA}^{-1} & \mathbf{I}_p \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{M}/\mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{I}_k & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix}$ (Theorem 8.3.1).
- False. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is invertible, but $\mathbf{A} = \mathbf{O}_1$ is not.
- The lower-right block, in the position of $\mathbf{D}$ (Theorem 8.3.4 (a)).
- True. Theorem 8.3.8 holds over every field and needs no invertibility; its proof cancels the nonzero polynomial $\det(\mathbf{D} + x\mathbf{I}_n)$ in $F[x]$.

8.3.7.2 B. Practice

Exercise. B1: A Schur complement, a determinant, and a rank

Let

$$\mathbf{M} = \left(\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 1 & 1 & 3 & 1 \\ \hline 1 & 2 & 9 & 5 \\ 0 & 1 & 6 & 4 \end{array} \right) \in M_4(\mathbb{Q}).$$

Compute $\mathbf{M}/\mathbf{A}$ for the upper-left block $\mathbf{A}$. Hence find $\det \mathbf{M}$ and $\text{rank } \mathbf{M}$, and decide whether $\mathbf{M}$ is invertible.

Solution. $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ has $\det \mathbf{A} = 1$, so $\mathbf{A}^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ (Theorem 0.9.8). With $\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$, $\mathbf{C} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 9 & 5 \\ 6 & 4 \end{pmatrix}$:

$$\mathbf{C}\mathbf{A}^{-1} = \begin{pmatrix} -1 & 3 \\ -1 & 2 \end{pmatrix}, \quad \mathbf{C}\mathbf{A}^{-1}\mathbf{B} = \begin{pmatrix} 8 & 3 \\ 5 & 2 \end{pmatrix}, \quad \mathbf{M}/\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}.$$

By Theorem 8.3.2, $\det \mathbf{M} = \det \mathbf{A} \det(\mathbf{M}/\mathbf{A}) = 1 \cdot 0 = 0$. By Theorem 8.3.3, $\text{rank } \mathbf{M} = 2 + \text{rank}(\mathbf{M}/\mathbf{A}) = 2 + 1 = 3$, since $\mathbf{M}/\mathbf{A}$ is nonzero with equal rows. So $\mathbf{M}$ is not invertible (Theorem 7.5.3 (d)).

Exercise. B2: A rank-one update

Let $\mathbf{A} = \text{diag}(1, 2, 4) \in M_3(\mathbb{Q})$, $\mathbf{u} = (1, 1, 1)$, and $\mathbf{v} = (1, -1, 2)$. Use Corollary 8.3.6 to find $\det(\mathbf{A} + \mathbf{u}\mathbf{v}^\top)$ and $(\mathbf{A} + \mathbf{u}\mathbf{v}^\top)^{-1}$, and check one entry.

Solution. $\mathbf{A}^{-1} = \text{diag}(1, \frac{1}{2}, \frac{1}{4})$, so $\mathbf{A}^{-1}\mathbf{u} = (1, \frac{1}{2}, \frac{1}{4})$ and $\mathbf{v}^\top \mathbf{A}^{-1} = (1, -\frac{1}{2}, \frac{1}{2})$. Then $\mathbf{v}^\top \mathbf{A}^{-1}\mathbf{u} = 1 - \frac{1}{2} + \frac{1}{2} = 1$, so $1 + \mathbf{v}^\top \mathbf{A}^{-1}\mathbf{u} = 2 \neq 0$. By Corollary 8.3.6, $\mathbf{A} + \mathbf{u}\mathbf{v}^\top$ is invertible, $\det(\mathbf{A} + \mathbf{u}\mathbf{v}^\top) = 8 \cdot 2 = 16$, and

$$\begin{aligned} (\mathbf{A} + \mathbf{u}\mathbf{v}^\top)^{-1} &= \text{diag}(1, \tfrac{1}{2}, \tfrac{1}{4}) - \frac{1}{2} \begin{pmatrix} 1 \\ \frac{1}{2} \\ \frac{1}{4} \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & -\frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & -\frac{1}{16} & \frac{1}{16} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{5}{8} & -\frac{1}{8} \\ -\frac{1}{8} & \frac{1}{16} & \frac{3}{16} \end{pmatrix}. \end{aligned}$$

Check: $\mathbf{A} + \mathbf{u}\mathbf{v}^\top = \begin{pmatrix} 2 & -1 & 2 \\ 1 & 1 & 2 \\ 1 & -1 & 6 \end{pmatrix}$, and its row 1 times column 1 of the answer is $1 + \frac{1}{4} - \frac{1}{4} = 1$, times column 2 is $\frac{1}{2} - \frac{5}{8} + \frac{1}{8} = 0$.

Exercise. B3: Identity corners

Let $\mathbf{B} \in M_{k \times l}(F)$, $\mathbf{C} \in M_{l \times k}(F)$, and $\mathbf{M} = \begin{pmatrix} \mathbf{I}_k & \mathbf{B} \\ \mathbf{C} & \mathbf{I}_l \end{pmatrix}$.

- Show that $\mathbf{M}$ is invertible if and only if $\mathbf{S} = \mathbf{I}_l - \mathbf{CB}$ is, and write down $\mathbf{M}^{-1}$ in terms of $\mathbf{S}^{-1}$.
- Deduce the inverse of $\begin{pmatrix} \mathbf{I}_k & \mathbf{B} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix}$.
- For $\mathbf{B} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\mathbf{C} = \begin{pmatrix} 1 & 1 \end{pmatrix}$ over $\mathbb{Q}$, compute $\mathbf{M}^{-1}$.

Solution.

- The block $\mathbf{I}_k$ is invertible and $\mathbf{M}/\mathbf{I}_k = \mathbf{I}_l - \mathbf{C}\mathbf{I}_k^{-1}\mathbf{B} = \mathbf{I}_l - \mathbf{CB}$. By [Theorem 8.3.4](#) (a), $\mathbf{M}$ is invertible if and only if $\mathbf{I}_l - \mathbf{CB}$ is, and then, with $\mathbf{A}^{-1} = \mathbf{I}_k$,

$$\mathbf{M}^{-1} = \begin{pmatrix} \mathbf{I}_k + \mathbf{B}\mathbf{S}^{-1}\mathbf{C} & -\mathbf{B}\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\mathbf{C} & \mathbf{S}^{-1} \end{pmatrix}.$$

- With $\mathbf{C} = \mathbf{O}$, $\mathbf{S} = \mathbf{I}_l - \mathbf{CB} = \mathbf{I}_l$ and $\mathbf{S}^{-1} = \mathbf{I}_l$, so the inverse is $\begin{pmatrix} \mathbf{I}_k & -\mathbf{B} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix}$, in agreement with [Theorem 8.1.5](#).

- $\mathbf{CB} = (3)$, so $\mathbf{S} = \mathbf{I}_1 - \mathbf{CB} = (-2)$ and $\mathbf{S}^{-1} = (-\frac{1}{2})$. Then $\mathbf{B}\mathbf{S}^{-1}\mathbf{C} = -\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$, $-\mathbf{B}\mathbf{S}^{-1} = \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}$, and $-\mathbf{S}^{-1}\mathbf{C} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix}$, so

$$\mathbf{M}^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -1 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}.$$

Check: row 3 of $\mathbf{M}$ times column 3 of $\mathbf{M}^{-1}$ is $\frac{1}{2} + 1 - \frac{1}{2} = 1$, and times column 1 is $\frac{1}{2} - 1 + \frac{1}{2} = 0$.

8.3.7.3 C. Going deeper

Exercise. C1: Sylvester's determinant identity

Let $\mathbf{A} \in M_{m \times n}(F)$ and $\mathbf{B} \in M_{n \times m}(F)$ with $m, n \geq 1$.

- (a) Compute both Schur complements of the identity blocks in $\mathbf{M} = \begin{pmatrix} \mathbf{I}_m & -\mathbf{A} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix}$, and deduce

$$\det(\mathbf{I}_m + \mathbf{A}\mathbf{B}) = \det(\mathbf{I}_n + \mathbf{B}\mathbf{A}).$$

- (b) Deduce that for every nonzero $c \in F$, $\det(c\mathbf{I}_m - \mathbf{A}\mathbf{B}) = c^{m-n} \det(c\mathbf{I}_n - \mathbf{B}\mathbf{A})$.
 (c) Deduce Theorem 7.8.5 from (a).

Solution.

- (a) Both identity blocks are invertible. By [Definition 8.3.1](#),

$$\mathbf{M}/\mathbf{I}_m = \mathbf{I}_n - \mathbf{B}\mathbf{I}_m^{-1}(-\mathbf{A}) = \mathbf{I}_n + \mathbf{B}\mathbf{A},$$

$$\mathbf{M}/\mathbf{I}_n = \mathbf{I}_m - (-\mathbf{A})\mathbf{I}_n^{-1}\mathbf{B} = \mathbf{I}_m + \mathbf{A}\mathbf{B}.$$

By [Theorem 8.3.2](#) (a) and (b), $\det \mathbf{M} = \det \mathbf{I}_m \det(\mathbf{I}_n + \mathbf{B}\mathbf{A})$ and $\det \mathbf{M} = \det \mathbf{I}_n \det(\mathbf{I}_m + \mathbf{A}\mathbf{B})$. Hence $\det(\mathbf{I}_m + \mathbf{A}\mathbf{B}) = \det(\mathbf{I}_n + \mathbf{B}\mathbf{A})$.

- (b) Let $\mathbf{A}' = -c^{-1}\mathbf{A}$. Then $c\mathbf{I}_m - \mathbf{A}\mathbf{B} = c(\mathbf{I}_m + \mathbf{A}'\mathbf{B})$ and $c\mathbf{I}_n - \mathbf{B}\mathbf{A} = c(\mathbf{I}_n + \mathbf{B}\mathbf{A}')$. Multiplying an $m \times m$ matrix by c multiplies each of its m rows by c , so it multiplies the determinant by c^m (Theorem 7.5.1). By (a) applied to $\mathbf{A}'$ and $\mathbf{B}$,

$$\begin{aligned} \det(c\mathbf{I}_m - \mathbf{A}\mathbf{B}) &= c^m \det(\mathbf{I}_m + \mathbf{A}'\mathbf{B}) = c^m \det(\mathbf{I}_n + \mathbf{B}\mathbf{A}') \\ &= c^m c^{-n} \det(c\mathbf{I}_n - \mathbf{B}\mathbf{A}) = c^{m-n} \det(c\mathbf{I}_n - \mathbf{B}\mathbf{A}). \end{aligned}$$

- (c) Let $\mathbf{u}, \mathbf{v} \in F^n$, and apply (a) with the sizes m, n there replaced by $n, 1$: $\mathbf{A} = \mathbf{u} \in M_{n \times 1}(F)$ and $\mathbf{B} = \mathbf{v}^\top \in M_{1 \times n}(F)$. Then $\det(\mathbf{I}_n + \mathbf{u}\mathbf{v}^\top) = \det(\mathbf{I}_1 + \mathbf{v}^\top \mathbf{u}) = 1 + \mathbf{v}^\top \mathbf{u}$, because the determinant of a 1×1 matrix is its entry.

Exercise. C2: Pivots as Schur complements

Let $\mathbf{A} \in M_n(F)$, and for $1 \leq k \leq n$ let $\mathbf{A}_k$ be its leading principal submatrix (the top-left $k \times k$ corner), as in Exercise (C1: Leading principal submatrices).

- (a) Suppose $\mathbf{A}_{k-1}$ is invertible, where $2 \leq k \leq n$, and write $\mathbf{A}_k = \begin{pmatrix} \mathbf{A}_{k-1} & \mathbf{b} \\ \mathbf{c}^\top & d \end{pmatrix}$. Show that $s_k := \mathbf{A}_k / \mathbf{A}_{k-1}$ is the scalar $d - \mathbf{c}^\top \mathbf{A}_{k-1}^{-1} \mathbf{b}$, and that $\det \mathbf{A}_k = \det \mathbf{A}_{k-1} \cdot s_k$.
 (b) Suppose $\mathbf{A}_1, \dots, \mathbf{A}_n$ are all invertible, and put $s_1 = a_{11}$. Prove that $\det \mathbf{A}_k = s_1 s_2 \cdots s_k$ for every k , and that every s_k is nonzero.
 (c) Keep the hypothesis of (b). In the proof of Exercise (C1: Leading principal submatrices) (c), the last pivot is $z = d - \mathbf{x}^\top \mathbf{y}$ with $\mathbf{y} = \mathbf{L}'^{-1} \mathbf{b}$ and $\mathbf{x}^\top = \mathbf{c}^\top \mathbf{U}'^{-1}$. Show that $z = \mathbf{A}_n / \mathbf{A}_{n-1}$. Deduce that the diagonal entries of $\mathbf{U}$ in the LU factorization are $u_{kk} = \det \mathbf{A}_k / \det \mathbf{A}_{k-1}$, with $\det \mathbf{A}_0 := 1$.

- (d) Check (c) for $\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 3 & 3 \\ 8 & 7 & 9 \end{pmatrix}$ over $\mathbb{Q}$.

Solution.

- (a) With the partition $k-1, 1$, Definition 8.3.1 gives the 1×1 matrix $\mathbf{A}_k/\mathbf{A}_{k-1} = d - \mathbf{c}^\top \mathbf{A}_{k-1}^{-1} \mathbf{b}$, which we identify with its entry. By Theorem 8.3.2 (a), $\det \mathbf{A}_k = \det \mathbf{A}_{k-1} \det(s_k) = \det \mathbf{A}_{k-1} \cdot s_k$.
- (b) By induction on k . For $k=1$, $\det \mathbf{A}_1 = a_{11} = s_1$. If $\det \mathbf{A}_{k-1} = s_1 \cdots s_{k-1}$, then by (a), which applies since $\mathbf{A}_{k-1}$ is invertible, $\det \mathbf{A}_k = s_1 \cdots s_{k-1} s_k$. Since $\mathbf{A}_k$ is invertible, $\det \mathbf{A}_k \neq 0$ (Theorem 7.5.2), so no factor s_k is zero (Theorem 0.7.1 (d)).
- (c) $\mathbf{A}_{n-1} = \mathbf{L}'\mathbf{U}'$ with both factors invertible, so $\mathbf{A}_{n-1}^{-1} = \mathbf{U}'^{-1}\mathbf{L}'^{-1}$ (Theorem 0.9.7 (3)) and $\mathbf{x}^\top \mathbf{y} = \mathbf{c}^\top \mathbf{U}'^{-1}\mathbf{L}'^{-1} \mathbf{b} = \mathbf{c}^\top \mathbf{A}_{n-1}^{-1} \mathbf{b}$. Hence $z = d - \mathbf{c}^\top \mathbf{A}_{n-1}^{-1} \mathbf{b} = \mathbf{A}_n/\mathbf{A}_{n-1} = s_n$ by (a). The same construction applied to $\mathbf{A}_k$ (whose LU factors are the top-left corners $\mathbf{L}_k, \mathbf{U}_k$ of $\mathbf{L}, \mathbf{U}$, by Exercise (C1: Leading principal submatrices) (a) and Theorem 3.6.5) shows $u_{kk} = s_k$ for every k . By (a) and (b), $s_k = \det \mathbf{A}_k / \det \mathbf{A}_{k-1}$, with $s_1 = \det \mathbf{A}_1 / 1$.
- (d) $\det \mathbf{A}_1 = 2$, $\det \mathbf{A}_2 = 6 - 4 = 2$, and expanding, $\det \mathbf{A}_3 = 2(27 - 21) - 1(36 - 24) + 1(28 - 24) = 12 - 12 + 4 = 4$. The predicted pivots are 2, $2/2 = 1$, $4/2 = 2$. Elimination gives $\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix}$ and $\mathbf{U} = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$: subtract 2 and 4 times row 1 to get rows $(0, 1, 1)$ and $(0, 3, 5)$, then 3 times the new row 2 to get $(0, 0, 2)$. The diagonal of $\mathbf{U}$ is 2, 1, 2, as predicted.

Exercise. C3: When A and C commute

Let $n \geq 1$ and $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D} \in M_n(F)$ with $\mathbf{AC} = \mathbf{CA}$, and $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$.

- (a) Assume $\mathbf{A}$ is invertible. Use Theorem 8.3.2 to prove $\det \mathbf{M} = \det(\mathbf{AD} - \mathbf{CB})$.
- (b) Remove the assumption that $\mathbf{A}$ is invertible, over any field, by adapting the proof of Theorem 8.3.8 with $\mathbf{A}_x = \mathbf{A} + x\mathbf{I}_n$.
- (c) For $\mathbf{A} = \mathbf{I}_2$, $\mathbf{B} = \mathbf{E}_{12}$, $\mathbf{C} = \mathbf{E}_{21}$, $\mathbf{D} = \mathbf{E}_{11}$ in $M_2(\mathbb{Q})$, check that $\mathbf{AC} = \mathbf{CA}$, and compare $\det \mathbf{M}$, $\det(\mathbf{AD} - \mathbf{CB})$, and $\det(\mathbf{AD} - \mathbf{BC})$.

Hint for (b): multiply $\mathbf{M}_x$ on the left by a block matrix that clears the lower-left block.

Solution.

- (a) By Theorem 8.3.2 (a) and Theorem 7.5.4,

$$\begin{aligned} \det \mathbf{M} &= \det \mathbf{A} \det(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}) \\ &= \det(\mathbf{A}(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})) = \det(\mathbf{AD} - \mathbf{ACA}^{-1}\mathbf{B}). \end{aligned}$$

Since $\mathbf{AC} = \mathbf{CA}$, $\mathbf{ACA}^{-1} = \mathbf{CAA}^{-1} = \mathbf{C}$, so $\det \mathbf{M} = \det(\mathbf{AD} - \mathbf{CB})$.

- (b) Let $\mathbf{A}_x = \mathbf{A} + x\mathbf{I}_n$ over $F[x]$. Then $\mathbf{A}_x \mathbf{C} = \mathbf{AC} + x\mathbf{C} = \mathbf{CA} + x\mathbf{C} = \mathbf{CA}_x$. Let

$\mathbf{M}_x = \begin{pmatrix} \mathbf{A}_x & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ and $\mathbf{L} = \begin{pmatrix} \mathbf{I}_n & \mathbf{O} \\ -\mathbf{C} & \mathbf{A}_x \end{pmatrix}$. By Theorem 8.1.1 over $F[x]$,

$$\mathbf{L}\mathbf{M}_x = \begin{pmatrix} \mathbf{A}_x & \mathbf{B} \\ -\mathbf{C}\mathbf{A}_x + \mathbf{A}_x\mathbf{C} & -\mathbf{C}\mathbf{B} + \mathbf{A}_x\mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_x & \mathbf{B} \\ \mathbf{O} & \mathbf{A}_x\mathbf{D} - \mathbf{C}\mathbf{B} \end{pmatrix}.$$

By Theorem 7.5.4 and Theorem 7.8.4 over $F[x]$ (Chapter 7 §5 and §8), $\det \mathbf{A}_x \det \mathbf{M}_x = \det \mathbf{A}_x \det(\mathbf{A}_x\mathbf{D} - \mathbf{C}\mathbf{B})$, using $\det \mathbf{L} = \det \mathbf{I}_n \det \mathbf{A}_x$. As in the proof of [Theorem 8.3.8](#), $\det \mathbf{A}_x = p_{-\mathbf{A}}(x)$ is monic, hence nonzero, and Corollary 0.8.4 (2) gives $\det \mathbf{M}_x = \det(\mathbf{A}_x\mathbf{D} - \mathbf{C}\mathbf{B})$. Evaluating at $x = 0$ with [Lemma 8.3.7](#) gives $\det \mathbf{M} = \det(\mathbf{A}\mathbf{D} - \mathbf{C}\mathbf{B})$.

- (c) $\mathbf{A}\mathbf{C} = \mathbf{C} = \mathbf{C}\mathbf{A}$, since $\mathbf{A} = \mathbf{I}_2$. As computed in Section 1, $\det \mathbf{M} = -1$. Also $\mathbf{A}\mathbf{D} - \mathbf{C}\mathbf{B} = \mathbf{E}_{11} - \mathbf{E}_{21}\mathbf{E}_{12} = \mathbf{E}_{11} - \mathbf{E}_{22} = \text{diag}(1, -1)$, with determinant -1 , in agreement with (b). But $\mathbf{A}\mathbf{D} - \mathbf{B}\mathbf{C} = \mathbf{E}_{11} - \mathbf{E}_{12}\mathbf{E}_{21} = \mathbf{E}_{11} - \mathbf{E}_{11} = \mathbf{O}$, with determinant 0. So with $\mathbf{A}\mathbf{C} = \mathbf{C}\mathbf{A}$ the product must be taken as $\mathbf{C}\mathbf{B}$, not $\mathbf{B}\mathbf{C}$; here $\mathbf{C}\mathbf{D} = \mathbf{E}_{21} \neq \mathbf{O} = \mathbf{D}\mathbf{C}$, so [Theorem 8.3.8](#) does not apply.