

8.2 Rank and Block Matrices

Rank is the most robust number attached to a matrix, and Chapters 3 and 4 found several inequalities for it: the rank of a product is at most the rank of each factor, and Sylvester's inequality bounds it from below. Those proofs worked with subspaces. This section proves rank inequalities by a different and very mechanical method. We place the matrices we care about as blocks of one bigger matrix, and then change that matrix by block row and column operations, which never change the rank, until its rank can be read off. The method gives short proofs of the Frobenius inequality and of identities such as $\text{rank}(\mathbf{I}_m - \mathbf{A}\mathbf{B}) + n = \text{rank}(\mathbf{I}_n - \mathbf{B}\mathbf{A}) + m$, which are awkward to reach with subspaces alone.

8.2.1 Three basic inequalities

We start with what the shape of a block matrix alone says about its rank. Write $\begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix}$ for the matrix obtained by placing $\mathbf{B}$ to the right of $\mathbf{A}$ (same number of rows), and $\begin{pmatrix} \mathbf{A} \\ \mathbf{C} \end{pmatrix}$ for $\mathbf{C}$ below $\mathbf{A}$ (same number of columns). In these results the blocks need **not** be square.

Theorem 8.2.1. Rank Inequalities for Block Matrices

Let $\mathbf{A} \in M_{m \times k}(F)$, $\mathbf{B} \in M_{m \times l}(F)$, $\mathbf{C} \in M_{p \times k}(F)$, and $\mathbf{D} \in M_{p \times l}(F)$.

- (a) $\max(\text{rank } \mathbf{A}, \text{rank } \mathbf{B}) \leq \text{rank} \begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix} \leq \text{rank } \mathbf{A} + \text{rank } \mathbf{B}$, and $\max(\text{rank } \mathbf{A}, \text{rank } \mathbf{C}) \leq \text{rank} \begin{pmatrix} \mathbf{A} \\ \mathbf{C} \end{pmatrix} \leq \text{rank } \mathbf{A} + \text{rank } \mathbf{C}$.
- (b) $\text{rank} \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} = \text{rank } \mathbf{A} + \text{rank } \mathbf{D}$.
- (c) $\text{rank} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} \geq \text{rank } \mathbf{A} + \text{rank } \mathbf{D}$ and $\text{rank} \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \geq \text{rank } \mathbf{A} + \text{rank } \mathbf{D}$.

Idea. Rank is the dimension of the column space, so look at columns. In (a) the columns of $\begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix}$ are those of $\mathbf{A}$ together with those of $\mathbf{B}$, and dimensions of sums are at most sums of dimensions. In (b) the columns coming from $\mathbf{A}$ and from $\mathbf{D}$ live in two subspaces that meet only in $\mathbf{0}$. In (c) we cannot use the column space of $\mathbf{B}$, which is unknown, so instead we pick independent columns of $\mathbf{A}$ and of $\mathbf{D}$ and show that the columns of the big matrix standing above them are still independent: the zero block lets us peel off the $\mathbf{D}$ -part first.

Proof. For $\mathbf{y} \in F^m$ and $\mathbf{z} \in F^p$, write $\begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} \in F^{m+p}$ for the stacked vector.

- (a) The columns of $\begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix}$ are the columns of $\mathbf{A}$ followed by the columns of $\mathbf{B}$, so its column space is $\text{col}(\mathbf{A}) + \text{col}(\mathbf{B})$ by Proposition 1.7.2. It contains $\text{col}(\mathbf{A})$ and $\text{col}(\mathbf{B})$, which gives the lower bound by Theorem 1.6.7, and by Theorem 1.7.3

$$\text{rank} \begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix} = \text{rank } \mathbf{A} + \text{rank } \mathbf{B} - \dim(\text{col } \mathbf{A} \cap \text{col } \mathbf{B}) \leq \text{rank } \mathbf{A} + \text{rank } \mathbf{B}.$$

For the stacked matrix, Proposition 8.1.2 gives $\begin{pmatrix} \mathbf{A} \\ \mathbf{C} \end{pmatrix}^\top = \begin{pmatrix} \mathbf{A}^\top & \mathbf{C}^\top \end{pmatrix}$, and a matrix and its transpose have the same rank (Theorem 3.5.5). Apply the first part to $\mathbf{A}^\top$ and $\mathbf{C}^\top$.

- (b) Let $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{pmatrix}$. Its first k columns are $\begin{pmatrix} \mathbf{a}_j \\ \mathbf{0} \end{pmatrix}$ and its last l columns are $\begin{pmatrix} \mathbf{0} \\ \mathbf{d}_q \end{pmatrix}$, where $\mathbf{a}_j, \mathbf{d}_q$ are the columns of $\mathbf{A}, \mathbf{D}$. So $\text{col}(\mathbf{M}) = Y + Z$ with $Y = \left\{ \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix} : \mathbf{y} \in \text{col } \mathbf{A} \right\}$ and $Z = \left\{ \begin{pmatrix} \mathbf{0} \\ \mathbf{z} \end{pmatrix} : \mathbf{z} \in \text{col } \mathbf{D} \right\}$. A vector in $Y \cap Z$ has both parts zero, so $Y \cap Z = \{\mathbf{0}\}$. The map $\mathbf{y} \mapsto \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix}$ is linear and injective from $\text{col } \mathbf{A}$ onto Y , so $\dim Y = \text{rank } \mathbf{A}$ (Theorem 2.5.4); likewise $\dim Z = \text{rank } \mathbf{D}$. By Theorem 1.7.3, $\text{rank } \mathbf{M} = \text{rank } \mathbf{A} + \text{rank } \mathbf{D} - 0$.
- (c) Let $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix}$, $r = \text{rank } \mathbf{A}$, and $s = \text{rank } \mathbf{D}$. By Theorem 3.5.2 there are columns $\mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_r}$ of $\mathbf{A}$ that are linearly independent, and columns $\mathbf{d}_{q_1}, \dots, \mathbf{d}_{q_s}$ of $\mathbf{D}$ that are linearly independent. The corresponding columns of $\mathbf{M}$ are $\begin{pmatrix} \mathbf{a}_{j_i} \\ \mathbf{0} \end{pmatrix}$ and $\begin{pmatrix} \mathbf{b}_{q_i} \\ \mathbf{d}_{q_i} \end{pmatrix}$, where $\mathbf{b}_q$ is column q of $\mathbf{B}$. Let

$$\sum_{i=1}^r \alpha_i \begin{pmatrix} \mathbf{a}_{j_i} \\ \mathbf{0} \end{pmatrix} + \sum_{i=1}^s \beta_i \begin{pmatrix} \mathbf{b}_{q_i} \\ \mathbf{d}_{q_i} \end{pmatrix} = \mathbf{0}.$$

The last p entries give $\sum_i \beta_i \mathbf{d}_{q_i} = \mathbf{0}$, so all $\beta_i = 0$. The first m entries then give $\sum_i \alpha_i \mathbf{a}_{j_i} = \mathbf{0}$, so all $\alpha_i = 0$. Hence $\text{col}(\mathbf{M})$ contains $r + s$ linearly independent vectors, and $\text{rank } \mathbf{M} \geq r + s$ by Theorem 1.6.3. For the lower triangular shape, $\begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^\top = \begin{pmatrix} \mathbf{A}^\top & \mathbf{C}^\top \\ \mathbf{O} & \mathbf{D}^\top \end{pmatrix}$ by Proposition 8.1.2; apply the first case and Theorem 3.5.5.

■

All three inequalities, the two in (a) and the one in (c), can be strict, and small witnesses show it. For $\mathbf{A} = \mathbf{B} = (1)$, $\text{rank} \begin{pmatrix} 1 & 1 \end{pmatrix} = 1 < 2$: the two column spaces overlap. For $\mathbf{A} = \mathbf{e}_1$ and $\mathbf{B} = \mathbf{e}_2$ in F^2 , $\text{rank} \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 \end{pmatrix} = \text{rank } \mathbf{I}_2 = 2 > \max(1, 1)$: each block adds something the other lacks. For $\mathbf{A} = \mathbf{D} = \mathbf{O}_1$ and $\mathbf{B} = (1)$, $\text{rank} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1 > 0 + 0$: the off-diagonal block contributes rank of its own. The proof of (a) says exactly when the upper bound is attained: $\text{rank} \begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix} = \text{rank } \mathbf{A} + \text{rank } \mathbf{B}$ if and only if $\text{col } \mathbf{A} \cap \text{col } \mathbf{B} = \{\mathbf{0}\}$.

Warning. The ranks of the four blocks do not determine the rank of the whole.

Over $\mathbb{Q}$, $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ have all four 1×1 blocks of rank 1, but ranks 1 and 2. Only special shapes, like the zero blocks in (b) and (c), turn block ranks into information about the whole matrix.

Quick check. Without computing, find $\text{rank} \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 2 & 4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 5 \end{pmatrix}$.

Solution. With row partition 2,2 and column partition 2,3, the matrix is $\begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{pmatrix}$ with $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ of rank 1 (the second row is twice the first) and $\mathbf{D} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 5 \end{pmatrix}$ of rank 2 (the first two columns are $\mathbf{e}_1, \mathbf{e}_2$). By [Theorem 8.2.1](#) (b), the rank is $1 + 2 = 3$.

8.2.2 Block elimination

Gaussian elimination changes a matrix by row operations, which are multiplications by invertible matrices on the left, and the rank never changes. The same is true when the “rows” are whole block rows and the multipliers are matrices. The matrices that do this are the block triangular matrices with identity blocks on the diagonal from [Theorem 8.1.5](#), together with block swaps.

Fix a partition of $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \in M_{(m+p) \times (k+l)}(F)$ with $\mathbf{A} \in M_{m \times k}(F)$. For matrices $\mathbf{X} \in M_{m \times p}(F)$ and $\mathbf{Y} \in M_{k \times l}(F)$, [Theorem 8.1.1](#) gives

$$\begin{pmatrix} \mathbf{I}_m & \mathbf{X} \\ \mathbf{O} & \mathbf{I}_p \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{A} + \mathbf{XC} & \mathbf{B} + \mathbf{XD} \\ \mathbf{C} & \mathbf{D} \end{pmatrix},$$

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{I}_k & \mathbf{Y} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{AY} + \mathbf{B} \\ \mathbf{C} & \mathbf{CY} + \mathbf{D} \end{pmatrix}.$$

In words: multiplying on the **left** adds $\mathbf{X}$ times block row 2 to block row 1, with $\mathbf{X}$ on the left of the row; multiplying on the **right** adds block column 1 times $\mathbf{Y}$ to block column 2, with $\mathbf{Y}$ on the right of the column. The lower triangular multipliers $\begin{pmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{X}' & \mathbf{I}_p \end{pmatrix}$ and $\begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ \mathbf{Y}' & \mathbf{I}_l \end{pmatrix}$ act the same way in the other direction. Sides matter: the multiplier sits on the side of the product where the matrix operation happens.

Theorem 8.2.2. Block Elimination Preserves Rank

Let $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ be as above.

(a) For all $\mathbf{X} \in M_{m \times p}(F)$, $\mathbf{X}' \in M_{p \times m}(F)$, $\mathbf{Y} \in M_{k \times l}(F)$, and $\mathbf{Y}' \in M_{l \times k}(F)$, the matrices

$$\begin{pmatrix} \mathbf{I}_m & \mathbf{X} \\ \mathbf{O} & \mathbf{I}_p \end{pmatrix}, \quad \begin{pmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{X}' & \mathbf{I}_p \end{pmatrix}, \quad \begin{pmatrix} \mathbf{I}_k & \mathbf{Y} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix}, \quad \begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ \mathbf{Y}' & \mathbf{I}_l \end{pmatrix}$$

are invertible, and so are the block swaps $\begin{pmatrix} \mathbf{O} & \mathbf{I}_p \\ \mathbf{I}_m & \mathbf{O} \end{pmatrix} \in M_{m+p}(F)$ and $\begin{pmatrix} \mathbf{O} & \mathbf{I}_k \\ \mathbf{I}_l & \mathbf{O} \end{pmatrix} \in M_{k+l}(F)$. Hence multiplying $\mathbf{M}$ on the left or right by any product of such matrices

does not change its rank. The swaps give $\begin{pmatrix} \mathbf{O} & \mathbf{I}_p \\ \mathbf{I}_m & \mathbf{O} \end{pmatrix} \mathbf{M} = \begin{pmatrix} \mathbf{C} & \mathbf{D} \\ \mathbf{A} & \mathbf{B} \end{pmatrix}$ and $\mathbf{M} \begin{pmatrix} \mathbf{O} & \mathbf{I}_k \\ \mathbf{I}_l & \mathbf{O} \end{pmatrix} = \begin{pmatrix} \mathbf{B} & \mathbf{A} \\ \mathbf{D} & \mathbf{C} \end{pmatrix}$.

(b) If $\text{col}(\mathbf{B}) \subseteq \text{col}(\mathbf{A})$, or if $\text{row}(\mathbf{B}) \subseteq \text{row}(\mathbf{D})$, then $\text{rank} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} = \text{rank } \mathbf{A} + \text{rank } \mathbf{D}$.

Proof.

- (a) By Theorem 8.1.5 (c), with identity blocks on the diagonal, $\begin{pmatrix} \mathbf{I}_m & \mathbf{X} \\ \mathbf{O} & \mathbf{I}_p \end{pmatrix}$ has inverse $\begin{pmatrix} \mathbf{I}_m & -\mathbf{X} \\ \mathbf{O} & \mathbf{I}_p \end{pmatrix}$; the lower triangular case follows by transposing (Proposition 8.1.2, Theorem 0.9.7). In the swap $\mathbf{S} = \begin{pmatrix} \mathbf{O} & \mathbf{I}_p \\ \mathbf{I}_m & \mathbf{O} \end{pmatrix}$ the row partition is p, m and the column partition is m, p . Let $\mathbf{S}' = \begin{pmatrix} \mathbf{O} & \mathbf{I}_m \\ \mathbf{I}_p & \mathbf{O} \end{pmatrix}$, with row partition m, p and column partition p, m . These are conformable, and Theorem 8.1.1 gives $\mathbf{S}'\mathbf{S} = \begin{pmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_p \end{pmatrix} = \mathbf{I}_{m+p}$, so $\mathbf{S}$ is invertible by Theorem 3.4.5; the other swap is the same with k, l in place of m, p . The two displayed products are Theorem 8.1.1. Rank is unchanged by invertible factors on either side (Theorem 3.5.9).
- (b) Suppose $\text{col}(\mathbf{B}) \subseteq \text{col}(\mathbf{A})$. Each column $\mathbf{b}_q$ of $\mathbf{B}$ is then $\mathbf{A}\mathbf{y}_q$ for some $\mathbf{y}_q \in F^k$ (Definition 3.5.1, Theorem 0.9.3). Let $\mathbf{Y} \in M_{k \times l}(F)$ have columns $\mathbf{y}_q$, so that $\mathbf{A}\mathbf{Y} = \mathbf{B}$ by Theorem 0.9.2. By the computation before the theorem,

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{I}_k & -\mathbf{Y} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix} = \begin{pmatrix} \mathbf{A} & -\mathbf{A}\mathbf{Y} + \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{pmatrix},$$

and (a) together with Theorem 8.2.1 (b) gives the claim. If $\text{row}(\mathbf{B}) \subseteq \text{row}(\mathbf{D})$, the same argument applied to rows gives $\mathbf{B} = \mathbf{X}\mathbf{D}$ for some $\mathbf{X} \in M_{m \times p}(F)$, and $\begin{pmatrix} \mathbf{I}_m & -\mathbf{X} \\ \mathbf{O} & \mathbf{I}_p \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{pmatrix}$.

■

Part (b) is the typical use of (a): a block that is a multiple of a neighboring block can be cleared, exactly as an entry below a pivot is cleared in Gaussian elimination. The zero block of Theorem 8.2.1 (c) is what lets the cleared matrix fall apart into a direct sum.

8.2.3 Frobenius's inequality by blocks

Recall Sylvester's rank inequality (Theorem 4.3.5): for $\mathbf{A} \in M_{m \times n}(F)$ and $\mathbf{B} \in M_{n \times p}(F)$, $\text{rank}(\mathbf{A}\mathbf{B}) \geq \text{rank } \mathbf{A} + \text{rank } \mathbf{B} - n$. In Exercise (C1: Frobenius's rank inequality) we met a stronger statement, Frobenius's inequality, and proved it by comparing intersections of subspaces with a null space. Here is a proof in which nothing is chosen: we write down one block matrix whose rank we know, and eliminate.

Theorem 8.2.3. Frobenius's Rank Inequality

Let $\mathbf{A} \in M_{m \times n}(F)$, $\mathbf{B} \in M_{n \times p}(F)$, and $\mathbf{C} \in M_{p \times q}(F)$. Then

$$\text{rank}(\mathbf{ABC}) + \text{rank} \mathbf{B} \geq \text{rank}(\mathbf{AB}) + \text{rank}(\mathbf{BC}).$$

Idea. The left side is the rank of $\begin{pmatrix} \mathbf{ABC} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \end{pmatrix}$, by Theorem 8.2.1 (b). The right side is a lower bound for any block triangular matrix with diagonal blocks $\mathbf{AB}$ and $\mathbf{BC}$, by part (c). So we want to turn the first matrix into the second by block elimination. Found backwards: $\mathbf{A}$ times the second block row puts $\mathbf{AB}$ into the top-right corner; then the second block column times $\mathbf{C}$ cancels $\mathbf{ABC}$ and puts $\mathbf{BC}$ into the bottom-left corner. A swap of the two block columns finishes the job.

Proof. Let $\mathbf{P} = \begin{pmatrix} \mathbf{ABC} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \end{pmatrix}$, with row partition m, n and column partition q, p . By Theorem 8.2.1 (b), $\text{rank} \mathbf{P} = \text{rank}(\mathbf{ABC}) + \text{rank} \mathbf{B}$. By Theorem 8.1.1,

$$\begin{aligned} \begin{pmatrix} \mathbf{I}_m & \mathbf{A} \\ \mathbf{O} & \mathbf{I}_n \end{pmatrix} \mathbf{P} &= \begin{pmatrix} \mathbf{ABC} & \mathbf{AB} \\ \mathbf{O} & \mathbf{B} \end{pmatrix}, \\ \begin{pmatrix} \mathbf{ABC} & \mathbf{AB} \\ \mathbf{O} & \mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{I}_q & \mathbf{O} \\ -\mathbf{C} & \mathbf{I}_p \end{pmatrix} &= \begin{pmatrix} \mathbf{ABC} - \mathbf{ABC} & \mathbf{AB} \\ -\mathbf{BC} & \mathbf{B} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{O} & \mathbf{AB} \\ -\mathbf{BC} & \mathbf{B} \end{pmatrix}. \end{aligned}$$

Multiplying on the right by the block swap $\begin{pmatrix} \mathbf{O} & \mathbf{I}_q \\ \mathbf{I}_p & \mathbf{O} \end{pmatrix}$ exchanges the two block columns and gives $\mathbf{Q} = \begin{pmatrix} \mathbf{AB} & \mathbf{O} \\ \mathbf{B} & -\mathbf{BC} \end{pmatrix}$. By Theorem 8.2.2 (a), $\text{rank} \mathbf{Q} = \text{rank} \mathbf{P}$. By Theorem 8.2.1 (c), $\text{rank} \mathbf{Q} \geq \text{rank}(\mathbf{AB}) + \text{rank}(-\mathbf{BC}) = \text{rank}(\mathbf{AB}) + \text{rank}(\mathbf{BC})$, where $\text{rank}(-\mathbf{BC}) = \text{rank}(\mathbf{BC})$ because $-\mathbf{BC}$ and $\mathbf{BC}$ have the same column space. Hence $\text{rank}(\mathbf{ABC}) + \text{rank} \mathbf{B} \geq \text{rank}(\mathbf{AB}) + \text{rank}(\mathbf{BC})$. ■

Taking $\mathbf{B} = \mathbf{I}_n$ gives $\text{rank}(\mathbf{AC}) + n \geq \text{rank} \mathbf{A} + \text{rank} \mathbf{C}$ for $\mathbf{A} \in M_{m \times n}(F)$ and $\mathbf{C} \in M_{n \times q}(F)$, which is Sylvester's inequality again. Taking $\mathbf{A} = \mathbf{I}_n$ or $\mathbf{C} = \mathbf{I}_p$ gives an equality and no information, as it should. The inequality says that multiplying by $\mathbf{A}$ on the left loses at least as much rank from $\mathbf{BC}$ as from $\mathbf{B}$: the rank lost is $\text{rank}(\mathbf{BC}) - \text{rank}(\mathbf{ABC}) \leq \text{rank} \mathbf{B} - \text{rank}(\mathbf{AB})$. A smaller column space has less to lose.

8.2.4 A worked example: $\mathbf{I} - \mathbf{AB}$ and $\mathbf{I} - \mathbf{BA}$

Products in the two orders, $\mathbf{AB}$ and $\mathbf{BA}$, are different matrices of possibly different sizes, and Chapter 4 showed that even their ranks can differ. Yet $\mathbf{I} - \mathbf{AB}$ and $\mathbf{I} - \mathbf{BA}$ are tightly linked. The block matrix that links them contains $\mathbf{I}_m$, $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{I}_n$, and each identity block can be

used as a pivot to clear its neighbors.

Example. Rank of $\mathbf{I} - \mathbf{AB}$ versus $\mathbf{I} - \mathbf{BA}$

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

- (a) Prove that $\text{rank}(\mathbf{I}_m - \mathbf{AB}) + n = \text{rank}(\mathbf{I}_n - \mathbf{BA}) + m$.
- (b) Deduce that for $m = n$, $\text{rank}(\mathbf{I}_n - \mathbf{AB}) = \text{rank}(\mathbf{I}_n - \mathbf{BA})$, and $\mathbf{I}_n - \mathbf{AB}$ is invertible if and only if $\mathbf{I}_n - \mathbf{BA}$ is.
- (c) Check (a) for $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$ over $\mathbb{Q}$.

Solution.

- (a) Let $\mathbf{M} = \begin{pmatrix} \mathbf{I}_m & \mathbf{A} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix} \in M_{m+n}(F)$, with partition m, n for rows and columns. We eliminate in two different ways.

Using $\mathbf{I}_m$ as the pivot. Subtract $\mathbf{B}$ times block row 1 from block row 2, then subtract block column 1 times $\mathbf{A}$ from block column 2. By Theorem 8.1.1,

$$\begin{pmatrix} \mathbf{I}_m & \mathbf{O} \\ -\mathbf{B} & \mathbf{I}_n \end{pmatrix} \mathbf{M} = \begin{pmatrix} \mathbf{I}_m & \mathbf{A} \\ \mathbf{O} & \mathbf{I}_n - \mathbf{BA} \end{pmatrix},$$

$$\begin{pmatrix} \mathbf{I}_m & \mathbf{A} \\ \mathbf{O} & \mathbf{I}_n - \mathbf{BA} \end{pmatrix} \begin{pmatrix} \mathbf{I}_m & -\mathbf{A} \\ \mathbf{O} & \mathbf{I}_n \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_n - \mathbf{BA} \end{pmatrix}.$$

By Theorem 8.2.2 (a) and Theorem 8.2.1 (b), $\text{rank } \mathbf{M} = m + \text{rank}(\mathbf{I}_n - \mathbf{BA})$.

Using $\mathbf{I}_n$ as the pivot. Subtract $\mathbf{A}$ times block row 2 from block row 1, then subtract block column 2 times $\mathbf{B}$ from block column 1:

$$\begin{pmatrix} \mathbf{I}_m & -\mathbf{A} \\ \mathbf{O} & \mathbf{I}_n \end{pmatrix} \mathbf{M} = \begin{pmatrix} \mathbf{I}_m - \mathbf{AB} & \mathbf{O} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix},$$

$$\begin{pmatrix} \mathbf{I}_m - \mathbf{AB} & \mathbf{O} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{I}_m & \mathbf{O} \\ -\mathbf{B} & \mathbf{I}_n \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m - \mathbf{AB} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_n \end{pmatrix}.$$

So $\text{rank } \mathbf{M} = \text{rank}(\mathbf{I}_m - \mathbf{AB}) + n$ in the same way. Comparing the two expressions for $\text{rank } \mathbf{M}$ proves (a).

- (b) For $m = n$ the two added terms are equal and cancel. A square matrix is invertible exactly when its rank is n (Theorem 7.5.3 (d)), so $\mathbf{I}_n - \mathbf{AB}$ is invertible if and only if $\mathbf{I}_n - \mathbf{BA}$ is.

- (c) Here $\mathbf{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{I}_2$, so $\text{rank}(\mathbf{I}_2 - \mathbf{AB}) = 0$, and

$$\mathbf{BA} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad \mathbf{I}_3 - \mathbf{BA} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix},$$

which has rank 1. Indeed $0 + 3 = 1 + 2$. So $\mathbf{AB} = \mathbf{I}_2$ while $\mathbf{BA} \neq \mathbf{I}_3$, and the identity accounts for the difference exactly: $\mathbf{I}_3 - \mathbf{BA}$ has rank $3 - 2 = 1$.

The trick is worth naming. **To compare two expressions, find one block matrix from which both can be obtained by elimination, and compute its rank twice.** Section 3 does the same with determinants and inverses, and the pair $\mathbf{I} - \mathbf{AB}$, $\mathbf{I} - \mathbf{BA}$ returns there as Sylvester's determinant identity.

8.2.5 Exercises

8.2.5.1 A. Check your understanding

Exercise. A1

- (a) State the three inequalities of [Theorem 8.2.1](#).
- (b) True or false: $\text{rank} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} = \text{rank } \mathbf{A} + \text{rank } \mathbf{D}$ for all blocks of compatible sizes. Justify your answer.
- (c) Which matrix do you multiply $\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ by, and on which side, to replace $\mathbf{C}$ by $\mathbf{C} - \mathbf{XA}$?
- (d) True or false: $\text{rank} \begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix} = \text{rank } \mathbf{A} + \text{rank } \mathbf{B}$ whenever $\mathbf{A}$ and $\mathbf{B}$ have the same number of rows. Justify your answer.
- (e) State Frobenius's inequality, and explain how Sylvester's inequality follows from it.

Solution.

- (a) $\max(\text{rank } \mathbf{A}, \text{rank } \mathbf{B}) \leq \text{rank} \begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix} \leq \text{rank } \mathbf{A} + \text{rank } \mathbf{B}$ (and the same for stacked blocks); $\text{rank} \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} = \text{rank } \mathbf{A} + \text{rank } \mathbf{D}$; and both block triangular shapes have rank at least $\text{rank } \mathbf{A} + \text{rank } \mathbf{D}$.
- (b) False. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ with 1×1 blocks has rank 1, but $\text{rank}(0) + \text{rank}(0) = 0$. Only $\geq$ holds in general.
- (c) Multiply on the left by $\begin{pmatrix} \mathbf{I} & \mathbf{O} \\ -\mathbf{X} & \mathbf{I} \end{pmatrix}$: block row 2 becomes $\begin{pmatrix} \mathbf{C} - \mathbf{XA} & \mathbf{D} - \mathbf{XB} \end{pmatrix}$ by Theorem 8.1.1. The rank does not change ([Theorem 8.2.2](#)).
- (d) False. $\begin{pmatrix} 1 & 1 \end{pmatrix}$ has rank 1, not $1 + 1$. Equality holds exactly when $\text{col } \mathbf{A} \cap \text{col } \mathbf{B} = \{\mathbf{0}\}$.
- (e) $\text{rank}(\mathbf{ABC}) + \text{rank } \mathbf{B} \geq \text{rank}(\mathbf{AB}) + \text{rank}(\mathbf{BC})$ for matrices of compatible sizes. With $\mathbf{B} = \mathbf{I}_n$ it reads $\text{rank}(\mathbf{AC}) + n \geq \text{rank } \mathbf{A} + \text{rank } \mathbf{C}$, which is Sylvester's inequality.

8.2.5.2 B. Practice

Exercise. B1: Ranks of block matrices

Let $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ and $\mathbf{D} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ in $M_2(\mathbb{Q})$. Find the rank of each 4×4 matrix, citing the result you use. In which case is the lower bound of [Theorem 8.2.1](#) (c) strict?

- (a) $\mathbf{M}_1 = \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{pmatrix}$.
- (b) $\mathbf{M}_2 = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix}$ with $\mathbf{B} = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$.
- (c) $\mathbf{M}_3 = \begin{pmatrix} \mathbf{A} & \mathbf{I}_2 \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$.

Solution. rank $\mathbf{A} = 1$, since the second row is twice the first and $\mathbf{A} \neq \mathbf{O}$; rank $\mathbf{D} = 2$, since $\det \mathbf{D} = 1 \neq 0$.

- (a) By [Theorem 8.2.1](#) (b), rank $\mathbf{M}_1 = 1 + 2 = 3$.
- (b) The columns of $\mathbf{B}$ are $(1, 2)$ and $(3, 6) = 3(1, 2)$, and $(1, 2)$ is the first column of $\mathbf{A}$. So $\text{col } \mathbf{B} \subseteq \text{col } \mathbf{A}$, and by [Theorem 8.2.2](#) (b), rank $\mathbf{M}_2 = 1 + 2 = 3$. (Explicitly $\mathbf{B} = \mathbf{A}\mathbf{Y}$ with $\mathbf{Y} = \begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}$.)
- (c) The lower bound is rank $\mathbf{A} + \text{rank } \mathbf{O} = 1$. But the last two columns of $\mathbf{M}_3$ are $\mathbf{e}_1, \mathbf{e}_2$, and every column lies in $\text{span}(\mathbf{e}_1, \mathbf{e}_2)$, so rank $\mathbf{M}_3 = 2$. This is the strict case: the block $\mathbf{I}_2$ is not a multiple of $\mathbf{A}$ or of the zero block.

Exercise. B2: Shifted products

Let $\mathbf{A} \in M_{m \times n}(F)$, $\mathbf{B} \in M_{n \times m}(F)$, and $c \in F$ with $c \neq 0$.

- (a) By adapting [Example \(Rank of \$\mathbf{I} - \mathbf{AB}\$ versus \$\mathbf{I} - \mathbf{BA}\$ \)](#), prove that $\text{rank}(c\mathbf{I}_m - \mathbf{AB}) + n = \text{rank}(c\mathbf{I}_n - \mathbf{BA}) + m$.
- (b) Check (a) over $\mathbb{Q}$ for $c = 2$, $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$, and $\mathbf{B} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.
- (c) Show by an example with $m = 1$, $n = 2$ that (a) can fail for $c = 0$.

Solution.

- (a) Since $c \neq 0$, put $\mathbf{A}' = c^{-1}\mathbf{A}$. Then $c\mathbf{I}_m - \mathbf{AB} = c(\mathbf{I}_m - \mathbf{A}'\mathbf{B})$ and $c\mathbf{I}_n - \mathbf{BA} = c(\mathbf{I}_n - \mathbf{BA}')$. Multiplying a matrix by the nonzero scalar c does not change its column space, hence not its rank. By [Example \(Rank of \$\mathbf{I} - \mathbf{AB}\$ versus \$\mathbf{I} - \mathbf{BA}\$ \)](#) (a) applied to $\mathbf{A}'$ and $\mathbf{B}$,

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

- (b) Here $m = 3$, $n = 2$, and

$$2\mathbf{I}_3 - \mathbf{AB} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad 2\mathbf{I}_2 - \mathbf{BA} = \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix},$$

since $\mathbf{AB} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ and $\mathbf{BA} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$. In the first matrix, row 1 equals $-(\text{row } 2) - (\text{row } 3)$, and rows 2 and 3 are not multiples of each other, so its rank is 2 (Theorem 3.5.5). The second has rank 1. Indeed $2 + 2 = 1 + 3$.

- (c) For $c = 0$ the identity would say $\text{rank}(\mathbf{AB}) + n = \text{rank}(\mathbf{BA}) + m$. Take $\mathbf{A} = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$: $\mathbf{AB} = (1)$ has rank 1, and $\mathbf{BA} = \mathbf{E}_{11} \in M_2(F)$ has rank 1, but $1 + 2 \neq 1 + 1$.

Exercise. B3: Idempotents by blocks

Let $\mathbf{A} \in M_n(F)$.

- (a) Starting from $\begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_n - \mathbf{A} \end{pmatrix}$, use block elimination to prove

$$\text{rank } \mathbf{A} + \text{rank}(\mathbf{I}_n - \mathbf{A}) = n + \text{rank}(\mathbf{A} - \mathbf{A}^2).$$

- (b) Deduce that $\mathbf{A}^2 = \mathbf{A}$ if and only if $\text{rank } \mathbf{A} + \text{rank}(\mathbf{I}_n - \mathbf{A}) = n$.
(c) Compare (b) with what Chapter 4 says about projections.

Hint: first add block row 1 to block row 2, then add block column 2 to block column 1.

Solution.

- (a) Write $\mathbf{I} = \mathbf{I}_n$. By Theorem 8.1.1, each step being multiplication by one of the matrices of Theorem 8.2.2 (a):

$$\begin{aligned} \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} - \mathbf{A} \end{pmatrix} &\xrightarrow{\text{row}_2 + \text{row}_1} \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{A} & \mathbf{I} - \mathbf{A} \end{pmatrix} \\ &\xrightarrow{\text{col}_1 + \text{col}_2} \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{I} & \mathbf{I} - \mathbf{A} \end{pmatrix} \\ &\xrightarrow{\text{row}_1 - \mathbf{A} \text{ row}_2} \begin{pmatrix} \mathbf{O} & \mathbf{A}^2 - \mathbf{A} \\ \mathbf{I} & \mathbf{I} - \mathbf{A} \end{pmatrix} \\ &\xrightarrow{\text{col}_2 - \text{col}_1(\mathbf{I} - \mathbf{A})} \begin{pmatrix} \mathbf{O} & \mathbf{A}^2 - \mathbf{A} \\ \mathbf{I} & \mathbf{O} \end{pmatrix}. \end{aligned}$$

The multipliers are, in order, $\begin{pmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{I} & \mathbf{I} \end{pmatrix}$ on the left, $\begin{pmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{I} & \mathbf{I} \end{pmatrix}$ on the right, $\begin{pmatrix} \mathbf{I} & -\mathbf{A} \\ \mathbf{O} & \mathbf{I} \end{pmatrix}$ on the left, and $\begin{pmatrix} \mathbf{I} & -(\mathbf{I} - \mathbf{A}) \\ \mathbf{O} & \mathbf{I} \end{pmatrix}$ on the right. In the third step the top row becomes $(\mathbf{A} - \mathbf{AI} \quad \mathbf{O} - \mathbf{A}(\mathbf{I} - \mathbf{A})) = (\mathbf{O} \quad \mathbf{A}^2 - \mathbf{A})$, and in the fourth the bottom-right block becomes $(\mathbf{I} - \mathbf{A}) - \mathbf{I}(\mathbf{I} - \mathbf{A}) = \mathbf{O}$. Swapping the block columns (Theorem 8.2.2 (a)) gives $\begin{pmatrix} \mathbf{A}^2 - \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{pmatrix}$, of rank $\text{rank}(\mathbf{A}^2 - \mathbf{A}) + n$ by Theorem 8.2.1 (b). The first

matrix has rank $\text{rank } \mathbf{A} + \text{rank}(\mathbf{I} - \mathbf{A})$, and rank is unchanged throughout. Since $\mathbf{A}^2 - \mathbf{A} = -(\mathbf{A} - \mathbf{A}^2)$ has the same rank as $\mathbf{A} - \mathbf{A}^2$, (a) follows.

- (b) By (a), $\text{rank } \mathbf{A} + \text{rank}(\mathbf{I} - \mathbf{A}) = n$ if and only if $\text{rank}(\mathbf{A} - \mathbf{A}^2) = 0$, that is, $\mathbf{A} - \mathbf{A}^2 = \mathbf{O}$.
- (c) If $\mathbf{A}^2 = \mathbf{A}$, then $T_{\mathbf{A}}$ is a projection (Definition 4.5.1), so $F^n = \text{im } T_{\mathbf{A}} \oplus \text{ker } T_{\mathbf{A}}$ by Theorem 4.5.1, with $\text{ker } T_{\mathbf{A}} = \text{col}(\mathbf{I} - \mathbf{A})$ (a vector $\mathbf{v}$ with $\mathbf{A}\mathbf{v} = \mathbf{O}$ equals $(\mathbf{I} - \mathbf{A})\mathbf{v}$, and conversely $\mathbf{A}(\mathbf{I} - \mathbf{A}) = \mathbf{A} - \mathbf{A}^2 = \mathbf{O}$). So $\text{rank } \mathbf{A} + \text{rank}(\mathbf{I} - \mathbf{A}) = \dim \text{im } T_{\mathbf{A}} + \dim \text{ker } T_{\mathbf{A}} = n$, which is the forward direction of (b). The block argument also gives the converse, and it never mentions subspaces.

8.2.5.3 C. Going deeper

Exercise. C1: The rank of $\mathbf{A} - \mathbf{ABA}$

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

- (a) Verify, by block multiplication, the following chain, in which each arrow is multiplication by an invertible block triangular matrix:

$$\begin{aligned} \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_n - \mathbf{BA} \end{pmatrix} &\rightarrow \begin{pmatrix} \mathbf{A} & \mathbf{A} \\ \mathbf{O} & \mathbf{I}_n - \mathbf{BA} \end{pmatrix} \rightarrow \begin{pmatrix} \mathbf{A} & \mathbf{A} \\ \mathbf{BA} & \mathbf{I}_n \end{pmatrix} \\ &\rightarrow \begin{pmatrix} \mathbf{A} - \mathbf{ABA} & \mathbf{A} \\ \mathbf{O} & \mathbf{I}_n \end{pmatrix} \rightarrow \begin{pmatrix} \mathbf{A} - \mathbf{ABA} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_n \end{pmatrix}. \end{aligned}$$

Name each multiplier and the side it acts on.

- (b) Deduce that $\text{rank}(\mathbf{A} - \mathbf{ABA}) = \text{rank } \mathbf{A} + \text{rank}(\mathbf{I}_n - \mathbf{BA}) - n$.
- (c) Suppose $\mathbf{ABA} = \mathbf{A}$. Show that $\mathbf{BA}$ is idempotent and that $\text{rank}(\mathbf{BA}) = \text{rank } \mathbf{A}$.

Hint for (c): use (b), then Exercise (B3: Idempotents by blocks).

Solution.

- (a) The partitions are m, n for rows and n, n for columns. By Theorem 8.1.1:

1. Right multiplication by $\begin{pmatrix} \mathbf{I}_n & \mathbf{I}_n \\ \mathbf{O} & \mathbf{I}_n \end{pmatrix}$ adds block column 1 to block column 2:

$$\begin{pmatrix} \mathbf{A} & \mathbf{A} + \mathbf{O} \\ \mathbf{O} & \mathbf{O} + (\mathbf{I}_n - \mathbf{BA}) \end{pmatrix}.$$

2. Left multiplication by $\begin{pmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix}$ adds $\mathbf{B}$ times block row 1 to block row 2: the new row is $\begin{pmatrix} \mathbf{BA} & \mathbf{BA} + \mathbf{I}_n - \mathbf{BA} \end{pmatrix} = \begin{pmatrix} \mathbf{BA} & \mathbf{I}_n \end{pmatrix}$.

3. Right multiplication by $\begin{pmatrix} \mathbf{I}_n & \mathbf{O} \\ -\mathbf{BA} & \mathbf{I}_n \end{pmatrix}$ subtracts block column 2 times $\mathbf{BA}$ from block column 1: the new column is $\begin{pmatrix} \mathbf{A} - \mathbf{ABA} \\ \mathbf{BA} - \mathbf{I}_n \mathbf{BA} \end{pmatrix} = \begin{pmatrix} \mathbf{A} - \mathbf{ABA} \\ \mathbf{O} \end{pmatrix}$.

4. Left multiplication by $\begin{pmatrix} \mathbf{I}_m & -\mathbf{A} \\ \mathbf{O} & \mathbf{I}_n \end{pmatrix}$ subtracts $\mathbf{A}$ times block row 2 from block row 1:

the new row is $(\mathbf{A} - \mathbf{A}\mathbf{B}\mathbf{A} - \mathbf{A} \cdot \mathbf{O} \quad \mathbf{A} - \mathbf{A}\mathbf{I}_n) = (\mathbf{A} - \mathbf{A}\mathbf{B}\mathbf{A} \quad \mathbf{O})$.

All four multipliers are invertible by [Theorem 8.2.2](#) (a).

- (b) By [Theorem 8.2.2](#) (a) the first and last matrices have the same rank, and by [Theorem 8.2.1](#) (b) these ranks are $\text{rank } \mathbf{A} + \text{rank}(\mathbf{I}_n - \mathbf{B}\mathbf{A})$ and $\text{rank}(\mathbf{A} - \mathbf{A}\mathbf{B}\mathbf{A}) + n$. Rearranging gives the formula.
- (c) $(\mathbf{B}\mathbf{A})^2 = \mathbf{B}(\mathbf{A}\mathbf{B}\mathbf{A}) = \mathbf{B}\mathbf{A}$, so $\mathbf{B}\mathbf{A}$ is idempotent. By (b), $0 = \text{rank}(\mathbf{A} - \mathbf{A}\mathbf{B}\mathbf{A}) = \text{rank } \mathbf{A} + \text{rank}(\mathbf{I}_n - \mathbf{B}\mathbf{A}) - n$, so $\text{rank}(\mathbf{I}_n - \mathbf{B}\mathbf{A}) = n - \text{rank } \mathbf{A}$. By [Exercise \(B3: Idempotents by blocks\)](#) (b) applied to the idempotent $\mathbf{B}\mathbf{A}$, $\text{rank}(\mathbf{B}\mathbf{A}) = n - \text{rank}(\mathbf{I}_n - \mathbf{B}\mathbf{A}) = \text{rank } \mathbf{A}$.

Exercise. C2: Involutions and the field

Let $\mathbf{A} \in M_n(F)$.

- (a) Suppose the characteristic of F is not 2. Starting from $\begin{pmatrix} \mathbf{A} - \mathbf{I}_n & \mathbf{O} \\ \mathbf{O} & \mathbf{A} + \mathbf{I}_n \end{pmatrix}$, prove that

$$\text{rank}(\mathbf{A} - \mathbf{I}_n) + \text{rank}(\mathbf{A} + \mathbf{I}_n) = n + \text{rank}(\mathbf{A}^2 - \mathbf{I}_n).$$

- (b) Deduce that if $\mathbf{A}^2 = \mathbf{I}_n$ (and $\text{char } F \neq 2$), then $\text{rank}(\mathbf{A} - \mathbf{I}_n) + \text{rank}(\mathbf{A} + \mathbf{I}_n) = n$.
- (c) Show that (b) fails over $\mathbb{F}_2$, and point to the step of (a) that breaks.

Hint for (a): add block column 2 to block column 1, and subtract block row 2 from block row 1; then $-2\mathbf{I}_n$ is an invertible pivot.

Solution.

- (a) Write $\mathbf{I} = \mathbf{I}_n$ and $\mathbf{P} = \mathbf{A} + \mathbf{I}$. Since $\text{char } F \neq 2$, $2 \neq 0$ in F , so $\frac{1}{2}$ exists and $-2\mathbf{I}$ is invertible. By [Theorem 8.1.1](#), with each multiplier from [Theorem 8.2.2](#) (a):

$$\begin{aligned} \begin{pmatrix} \mathbf{A} - \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{P} \end{pmatrix} &\xrightarrow{\text{col}_1 + \text{col}_2} \begin{pmatrix} \mathbf{A} - \mathbf{I} & \mathbf{O} \\ \mathbf{P} & \mathbf{P} \end{pmatrix} \xrightarrow{\text{row}_1 - \text{row}_2} \begin{pmatrix} -2\mathbf{I} & -\mathbf{P} \\ \mathbf{P} & \mathbf{P} \end{pmatrix} \\ &\xrightarrow{\text{col}_2 - \text{col}_1 \cdot \frac{1}{2}\mathbf{P}} \begin{pmatrix} -2\mathbf{I} & \mathbf{O} \\ \mathbf{P} & \mathbf{P} - \frac{1}{2}\mathbf{P}^2 \end{pmatrix} \\ &\xrightarrow{\text{row}_2 + \frac{1}{2}\mathbf{P} \text{ row}_1} \begin{pmatrix} -2\mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{P} - \frac{1}{2}\mathbf{P}^2 \end{pmatrix}. \end{aligned}$$

In the second step $(\mathbf{A} - \mathbf{I}) - \mathbf{P} = -2\mathbf{I}$; in the third, $-\mathbf{P} - (-2\mathbf{I})\frac{1}{2}\mathbf{P} = \mathbf{O}$ and $\mathbf{P} - \mathbf{P} \cdot \frac{1}{2}\mathbf{P} = \mathbf{P} - \frac{1}{2}\mathbf{P}^2$; in the fourth, $\mathbf{P} + \frac{1}{2}\mathbf{P}(-2\mathbf{I}) = \mathbf{O}$. Finally

$$\mathbf{P} - \frac{1}{2}\mathbf{P}^2 = \frac{1}{2}\mathbf{P}(2\mathbf{I} - \mathbf{P}) = \frac{1}{2}(\mathbf{A} + \mathbf{I})(\mathbf{I} - \mathbf{A}) = \frac{1}{2}(\mathbf{I} - \mathbf{A}^2),$$

where the product $(\mathbf{A} + \mathbf{I})(\mathbf{I} - \mathbf{A}) = \mathbf{I} - \mathbf{A} + \mathbf{A} - \mathbf{A}^2$ expands by distributivity. The first matrix has rank $\text{rank}(\mathbf{A} - \mathbf{I}) + \text{rank}(\mathbf{A} + \mathbf{I})$ and the last has rank $n + \text{rank}(\frac{1}{2}(\mathbf{I} - \mathbf{A}^2)) = n + \text{rank}(\mathbf{A}^2 - \mathbf{I})$ by [Theorem 8.2.1](#) (b), nonzero scalar multiples having the same column space. Rank is unchanged throughout.

- (b) If $\mathbf{A}^2 = \mathbf{I}_n$, then $\text{rank}(\mathbf{A}^2 - \mathbf{I}_n) = 0$.

- (c) Over $\mathbb{F}_2$, take $\mathbf{A} = \mathbf{I}_n$. Then $\mathbf{A}^2 = \mathbf{I}_n$, but $\mathbf{A} - \mathbf{I}_n = \mathbf{O}$ and $\mathbf{A} + \mathbf{I}_n = 2\mathbf{I}_n = \mathbf{O}$, so the left side is $0 \neq n$. The step that breaks is the pivot: $-2\mathbf{I}_n = \mathbf{O}$ is not invertible, and $\frac{1}{2}$ does not exist. (Chapter 2's decomposition of an involution into ± 1 parts, Theorem 4.5.5, needs the same hypothesis.)