5.4 Rank (in)equalities

Exercise 5.41 (Two zero blocks, rank)
(a) For any two matrices $\boldsymbol{A}$ and $\boldsymbol{D}$ (not necessarily square), show that

$$\operatorname{rk}\begin{pmatrix} \boldsymbol{A} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{D} \end{pmatrix} = \operatorname{rk}(\boldsymbol{A}) + \operatorname{rk}(\boldsymbol{D}).$$

(b) For any two matrices $\boldsymbol{B}$ and $\boldsymbol{C}$ (not necessarily square), show that

$$\operatorname{rk}\begin{pmatrix} \mathbf{O} & \boldsymbol{B} \\ \boldsymbol{C} & \mathbf{O} \end{pmatrix} = \operatorname{rk}(\boldsymbol{B}) + \operatorname{rk}(\boldsymbol{C}).$$

Solution
(a) The rank of a matrix is equal to the number of its linearly independent columns. Let

$$\boldsymbol{Z} := \begin{pmatrix} \boldsymbol{A} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{D} \end{pmatrix}, \quad \widetilde{\boldsymbol{A}} := \begin{pmatrix} \boldsymbol{A} \\ \mathbf{O} \end{pmatrix}, \quad \widetilde{\boldsymbol{D}} := \begin{pmatrix} \mathbf{O} \\ \boldsymbol{D} \end{pmatrix}.$$

Let $\widetilde{\boldsymbol{a}} := (\boldsymbol{a}', \mathbf{0}')'$ and $\widetilde{\boldsymbol{d}} := (\mathbf{0}', \boldsymbol{d}')'$ be two nonzero columns of $\widetilde{\boldsymbol{A}}$ and $\widetilde{\boldsymbol{D}}$, respectively. Then $\widetilde{\boldsymbol{a}}$ and $\widetilde{\boldsymbol{d}}$ are linearly independent, because if

$$\lambda_1 \widetilde{\boldsymbol{a}} + \lambda_2 \widetilde{\boldsymbol{d}} = \lambda_1 \begin{pmatrix} \boldsymbol{a} \\ \mathbf{0} \end{pmatrix} + \lambda_2 \begin{pmatrix} \mathbf{0} \\ \boldsymbol{d} \end{pmatrix} = \begin{pmatrix} \lambda_1 \boldsymbol{a} \\ \lambda_2 \boldsymbol{d} \end{pmatrix} = \mathbf{0},$$

then $\lambda_1 = \lambda_2 = 0$ (since $\widetilde{\boldsymbol{a}}$ and $\widetilde{\boldsymbol{d}}$ are nonzero). This implies that $\operatorname{rk}(\widetilde{\boldsymbol{A}} : \widetilde{\boldsymbol{D}}) = \operatorname{rk}(\widetilde{\boldsymbol{A}}) + \operatorname{rk}(\widetilde{\boldsymbol{D}})$ and hence that $\operatorname{rk}(\boldsymbol{Z}) = \operatorname{rk}(\boldsymbol{A}) + \operatorname{rk}(\boldsymbol{D})$.
(b) The rank does not change if we interchange columns. Hence,

$$\operatorname{rk}\begin{pmatrix} \mathbf{O} & \boldsymbol{B} \\ \boldsymbol{C} & \mathbf{O} \end{pmatrix} = \operatorname{rk}\begin{pmatrix} \boldsymbol{B} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{C} \end{pmatrix} = \operatorname{rk}(\boldsymbol{B}) + \operatorname{rk}(\boldsymbol{C}),$$

using (a).

Exercise 5.42 (One off-diagonal zero block, rank) Consider the matrices

$$\boldsymbol{Z}_1 := \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \mathbf{O} & \boldsymbol{D} \end{pmatrix} \quad \text{and} \quad \boldsymbol{Z}_2 := \begin{pmatrix} \boldsymbol{A} & \mathbf{O} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}.$$

Show that it is *not* true, in general, that $\operatorname{rk}(\boldsymbol{Z}_1) = \operatorname{rk}(\boldsymbol{A}) + \operatorname{rk}(\boldsymbol{D})$ or that $\operatorname{rk}(\boldsymbol{Z}_2) = \operatorname{rk}(\boldsymbol{A}) + \operatorname{rk}(\boldsymbol{D})$.

Solution
Take $\boldsymbol{A} = \mathbf{O}$ and $\boldsymbol{D} = \mathbf{O}$. Then $\operatorname{rk}(\boldsymbol{A}) = \operatorname{rk}(\boldsymbol{D}) = 0$, but $\operatorname{rk}(\boldsymbol{Z}_1) = \operatorname{rk}(\boldsymbol{B})$ and $\operatorname{rk}(\boldsymbol{Z}_2) = \operatorname{rk}(\boldsymbol{C})$, which are not zero, unless $\boldsymbol{B} = \mathbf{O}$ and $\boldsymbol{C} = \mathbf{O}$.

Exercise 5.43 (Nonsingular diagonal block, rank) Consider the matrices $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ of Exercise 5.42. If either $\boldsymbol{A}$ or $\boldsymbol{D}$ (or both) is nonsingular, show that

$$\operatorname{rk}(\boldsymbol{Z}_1) = \operatorname{rk}(\boldsymbol{Z}_2) = \operatorname{rk}(\boldsymbol{A}) + \operatorname{rk}(\boldsymbol{D}).$$

Is this condition necessary?

Solution
First, if $A = I_m$ and $D = I_n$, then both Z_1 and Z_2 are nonsingular (their determinant is 1 by Exercise 5.25). Now assume that $|A| \neq 0$. Then,

$$\begin{pmatrix} A & B \\ O & D \end{pmatrix} \begin{pmatrix} I_m & -A^{-1}B \\ O & I_q \end{pmatrix} = \begin{pmatrix} A & O \\ O & D \end{pmatrix} = \begin{pmatrix} I_m & O \\ -CA^{-1} & I_n \end{pmatrix} \begin{pmatrix} A & O \\ C & D \end{pmatrix}$$

and the result follows from Exercise 4.24. Similarly, if $|D| \neq 0$, we have

$$\begin{pmatrix} I_m & -BD^{-1} \\ O & I_n \end{pmatrix} \begin{pmatrix} A & B \\ O & D \end{pmatrix} = \begin{pmatrix} A & O \\ O & D \end{pmatrix} = \begin{pmatrix} A & O \\ C & D \end{pmatrix} \begin{pmatrix} I_p & O \\ -D^{-1}C & I_n \end{pmatrix}.$$

The condition is not necessary. For example, if $B = O$ and $C = O$, then $\operatorname{rk}(Z_1)$ and $\operatorname{rk}(Z_2)$ are both equal to $\operatorname{rk}(A) + \operatorname{rk}(D)$ whatever the ranks of A and D.

Exercise 5.44 (Nonsingular off-diagonal block, rank) Consider again the matrices Z_1 and Z_2 of Exercise 5.42. Show that

$$\operatorname{rk}(Z_1) = \operatorname{rk}(B) + \operatorname{rk}(DB^{-1}A)$$

if B is square and nonsingular, and

$$\operatorname{rk}(Z_2) = \operatorname{rk}(C) + \operatorname{rk}(AC^{-1}D)$$

if C is square and nonsingular.

Solution
The results follow from the equalities

$$\begin{pmatrix} I_m & O \\ -DB^{-1} & I_n \end{pmatrix} \begin{pmatrix} A & B \\ O & D \end{pmatrix} \begin{pmatrix} O & I_p \\ I_m & -B^{-1}A \end{pmatrix} = \begin{pmatrix} B & O \\ O & -DB^{-1}A \end{pmatrix}$$

and

$$\begin{pmatrix} O & I_n \\ I_m & -AC^{-1} \end{pmatrix} \begin{pmatrix} A & O \\ C & D \end{pmatrix} \begin{pmatrix} I_n & -C^{-1}D \\ O & I_q \end{pmatrix} = \begin{pmatrix} C & O \\ O & -AC^{-1}D \end{pmatrix}.$$

Exercise 5.45 (Rank inequalities, 1)
(a) Prove that

$$\operatorname{rk}\begin{pmatrix} A & B \\ O & D \end{pmatrix} \geq \operatorname{rk}(A) + \operatorname{rk}(D), \quad \operatorname{rk}\begin{pmatrix} A & O \\ C & D \end{pmatrix} \geq \operatorname{rk}(A) + \operatorname{rk}(D).$$

(b) Show that it is not true, in general, that

$$\operatorname{rk}\begin{pmatrix} A & B \\ C & D \end{pmatrix} \geq \operatorname{rk}(A) + \operatorname{rk}(D).$$

Solution
(a) Let

$$Z := \begin{pmatrix} A & B \\ O & D \end{pmatrix},$$

where the orders of the matrices are: A $(m \times p)$, B $(m \times q)$, and D $(n \times q)$. Suppose that $r := \operatorname{rk}(A) \leq p$ and that $s := \operatorname{rk}(D) \leq q$. Then A has r linearly independent columns, say $a_1, \dots, a_r$, and D has s linearly independent columns, say $d_1, \dots, d_s$. Let b_j denote the column of B directly above d_j in the matrix Z. Now consider the set of $r + s$ columns of Z,

$$\begin{pmatrix} a_1 \\ 0 \end{pmatrix}, \begin{pmatrix} a_2 \\ 0 \end{pmatrix}, \begin{pmatrix} a_r \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} b_1 \\ d_1 \end{pmatrix}, \begin{pmatrix} b_2 \\ d_2 \end{pmatrix}, \begin{pmatrix} b_s \\ d_s \end{pmatrix}.$$

We shall show that these $r + s$ columns are linearly independent. Suppose they are linearly dependent. Then there exist numbers $\alpha_1, \dots, \alpha_r$ and $\beta_1, \dots, \beta_s$, not all zero, such that

$$\sum_{i=1}^{r} \alpha_i \begin{pmatrix} a_i \\ 0 \end{pmatrix} + \sum_{j=1}^{s} \beta_j \begin{pmatrix} b_j \\ d_j \end{pmatrix} = 0.$$

This gives the two equations

$$\sum_{i=1}^{r} \alpha_i a_i + \sum_{j=1}^{s} \beta_j b_j = 0, \quad \sum_{j=1}^{s} \beta_j d_j = 0.$$

Since the $\{d_j\}$ are linearly independent, the second equation implies that $\beta_j = 0$ for all j. The first equation then reduces to $\sum_{i=1}^{r} \alpha_i a_i = 0$. Since the $\{a_i\}$ are linearly independent as well, all α_i are zero. We now have a contradiction. The matrix Z thus possesses (at least) $r + s$ linearly independent columns, so that $\operatorname{rk}(Z) \geq r + s = \operatorname{rk}(A) + \operatorname{rk}(D)$.

The second result can be proved analogously. Alternatively, it can be proved from the first result by considering the transpose:

$$\begin{aligned} \operatorname{rk} \begin{pmatrix} A & O \\ C & D \end{pmatrix} &= \operatorname{rk} \begin{pmatrix} A & O \\ C & D \end{pmatrix}' = \operatorname{rk} \begin{pmatrix} A' & C' \\ O & D' \end{pmatrix} \\ &\geq \operatorname{rk}(A') + \operatorname{rk}(D') = \operatorname{rk}(A) + \operatorname{rk}(D). \end{aligned}$$

(b) Consider

$$Z := \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} I_m & I_m \\ I_m & I_m \end{pmatrix}.$$

Then $\operatorname{rk}(A) = \operatorname{rk}(D) = \operatorname{rk}(Z) = m$, so that the inequality does not hold.

Exercise 5.46 (Rank inequalities, 2) Consider the matrices

$$Z_1 := \begin{pmatrix} A & B \\ C & O \end{pmatrix} \quad \text{and} \quad Z_2 := \begin{pmatrix} O & B \\ C & D \end{pmatrix}.$$

(a) If either B or C (or both) is nonsingular, then show that

$$\operatorname{rk}(Z_1) = \operatorname{rk}(Z_2) = \operatorname{rk}(B) + \operatorname{rk}(C).$$

(b) Show that

$$\operatorname{rk}(Z_1) = \operatorname{rk}(A) + \operatorname{rk}(CA^{-1}B)$$

if A is square and nonsingular, and

$$\operatorname{rk}(Z_2) = \operatorname{rk}(D) + \operatorname{rk}(BD^{-1}C)$$

if D is square and nonsingular.
(c) Show that

$$\operatorname{rk}(Z_1) \geq \operatorname{rk}(B) + \operatorname{rk}(C), \quad \operatorname{rk}(Z_2) \geq \operatorname{rk}(B) + \operatorname{rk}(C).$$

Solution
Since the rank does not change if we interchange columns, we have

$$\operatorname{rk}(Z_1) = \operatorname{rk}\begin{pmatrix} B & A \\ O & C \end{pmatrix}, \quad \operatorname{rk}(Z_2) = \operatorname{rk}\begin{pmatrix} B & O \\ D & C \end{pmatrix}.$$

Results (a)–(c) now follow from Exercises 5.43–5.45.

Exercise 5.47 (The inequalities of Frobenius and Sylvester)
(a) Use Exercise 5.46 to obtain the following famous inequality:

$$\operatorname{rk}(AB) + \operatorname{rk}(BC) \leq \operatorname{rk}(B) + \operatorname{rk}(ABC),$$

if the product ABC exists (*Frobenius*).
(b) From (a) obtain another famous inequality:

$$\operatorname{rk}(AB) \geq \operatorname{rk}(A) + \operatorname{rk}(B) - p$$

for any $m \times p$ matrix A and $p \times n$ matrix B (*Sylvester's law of nullity*).
(c) Show that $AB = O$ implies that $\operatorname{rk}(A) \leq p - \operatorname{rk}(B)$ for any $m \times p$ matrix A and $p \times n$ matrix B. (This generalizes Exercise 4.8.)

Solution
(a) Consider the identity

$$\begin{pmatrix} I_m & -A \\ O & I_n \end{pmatrix}\begin{pmatrix} O & AB \\ BC & B \end{pmatrix}\begin{pmatrix} I_q & O \\ -C & I_p \end{pmatrix} = \begin{pmatrix} -ABC & O \\ O & B \end{pmatrix}.$$

Of the four matrices, the first and third are nonsingular. Hence,

$$\operatorname{rk}\begin{pmatrix} O & AB \\ BC & B \end{pmatrix} = \operatorname{rk}(ABC) + \operatorname{rk}(B).$$

Also, by Exercise 5.46(c),

$$\operatorname{rk}\begin{pmatrix} \mathbf{O} & \boldsymbol{AB} \\ \boldsymbol{BC} & \boldsymbol{B} \end{pmatrix} \geq \operatorname{rk}(\boldsymbol{AB}) + \operatorname{rk}(\boldsymbol{BC}),$$

and the result follows.
(b) From Frobenius's inequality we obtain

$$\operatorname{rk}(\boldsymbol{AX}) + \operatorname{rk}(\boldsymbol{XB}) \leq \operatorname{rk}(\boldsymbol{X}) + \operatorname{rk}(\boldsymbol{AXB})$$

for any square matrix $\boldsymbol{X}$ of order p. Setting $\boldsymbol{X} = \boldsymbol{I}_p$ gives the result.
(c) Since $\boldsymbol{AB} = \mathbf{O}$, Sylvester's inequality gives $0 = \operatorname{rk}(\boldsymbol{AB}) \geq \operatorname{rk}(\boldsymbol{A}) + \operatorname{rk}(\boldsymbol{B}) - p$.

Exercise 5.48 (Rank of a partitioned matrix: main result) Let

$$\boldsymbol{Z} := \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}.$$

Show that

$$\operatorname{rk}(\boldsymbol{Z}) = \operatorname{rk}(\boldsymbol{A}) + \operatorname{rk}(\boldsymbol{D} - \boldsymbol{CA}^{-1}\boldsymbol{B}) \quad (\text{if } |\boldsymbol{A}| \neq 0)$$

and

$$\operatorname{rk}(\boldsymbol{Z}) = \operatorname{rk}(\boldsymbol{D}) + \operatorname{rk}(\boldsymbol{A} - \boldsymbol{BD}^{-1}\boldsymbol{C}) \quad (\text{if } |\boldsymbol{D}| \neq 0).$$

Solution
If $\boldsymbol{A}$ is nonsingular we can write

$$\begin{pmatrix} \boldsymbol{I}_m & \mathbf{O} \\ -\boldsymbol{CA}^{-1} & \boldsymbol{I}_n \end{pmatrix} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \begin{pmatrix} \boldsymbol{I}_m & -\boldsymbol{A}^{-1}\boldsymbol{B} \\ \mathbf{O} & \boldsymbol{I}_q \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{D} - \boldsymbol{CA}^{-1}\boldsymbol{B} \end{pmatrix}.$$

Similarly, if $\boldsymbol{D}$ is nonsingular, we can write

$$\begin{pmatrix} \boldsymbol{I}_m & -\boldsymbol{BD}^{-1} \\ \mathbf{O} & \boldsymbol{I}_n \end{pmatrix} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \begin{pmatrix} \boldsymbol{I}_p & \mathbf{O} \\ -\boldsymbol{D}^{-1}\boldsymbol{C} & \boldsymbol{I}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} - \boldsymbol{BD}^{-1}\boldsymbol{C} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{D} \end{pmatrix}.$$

Since for any matrix $\boldsymbol{Z}$, $\operatorname{rk}(\boldsymbol{Z}) = \operatorname{rk}(\boldsymbol{EZF})$ whenever $\boldsymbol{E}$ and $\boldsymbol{F}$ are nonsingular, the results follow.

Exercise 5.49 (Relationship between the ranks of $\boldsymbol{I}_m - \boldsymbol{BB}'$ and $\boldsymbol{I}_n - \boldsymbol{B}'\boldsymbol{B}$) Show that

$$\operatorname{rk}\begin{pmatrix} \boldsymbol{I}_m & \boldsymbol{B} \\ \boldsymbol{B}' & \boldsymbol{I}_n \end{pmatrix} = m + \operatorname{rk}(\boldsymbol{I}_n - \boldsymbol{B}'\boldsymbol{B}) = n + \operatorname{rk}(\boldsymbol{I}_m - \boldsymbol{BB}').$$

Solution
From Exercise 5.48 we obtain

$$\operatorname{rk}\begin{pmatrix} \boldsymbol{I}_m & \boldsymbol{B} \\ \boldsymbol{B}' & \boldsymbol{I}_n \end{pmatrix} = \operatorname{rk}(\boldsymbol{I}_m) + \operatorname{rk}(\boldsymbol{I}_n - \boldsymbol{B}'\boldsymbol{B}) = m + \operatorname{rk}(\boldsymbol{I}_n - \boldsymbol{B}'\boldsymbol{B})$$

and also

$$\operatorname{rk}\begin{pmatrix} I_m & B \\ B' & I_n \end{pmatrix} = \operatorname{rk}(I_n) + \operatorname{rk}(I_m - BB') = n + \operatorname{rk}(I_m - BB').$$

Exercise 5.50 (Relationship between the ranks of $I_m - BC$ and $I_n - CB$)
(a) Let B and C be square $n \times n$ matrices. Show that

$$\operatorname{rk}(I_n - BC) = \operatorname{rk}(I_n - CB).$$

(b) Now let B be an $m \times n$ matrix and C an $n \times m$ matrix. Extend the result under (a) by showing that

$$\operatorname{rk}(I_m - BC) = \operatorname{rk}(I_n - CB) + m - n.$$

Solution
(a) We have

$$\begin{pmatrix} I_n & -B \\ O & I_n \end{pmatrix} \begin{pmatrix} I_n & B \\ C & I_n \end{pmatrix} \begin{pmatrix} I_n & O \\ -C & I_n \end{pmatrix} = \begin{pmatrix} I_n - BC & O \\ O & I_n \end{pmatrix}$$

and

$$\begin{pmatrix} I_n & O \\ -C & I_n \end{pmatrix} \begin{pmatrix} I_n & B \\ C & I_n \end{pmatrix} \begin{pmatrix} I_n & -B \\ O & I_n \end{pmatrix} = \begin{pmatrix} I_n & O \\ O & I_n - CB \end{pmatrix}.$$

This proves (a) and shows in addition that

$$\operatorname{rk}(I_n - BC) = \operatorname{rk}(I_n - CB) = \operatorname{rk}\begin{pmatrix} I_n & B \\ C & I_n \end{pmatrix}.$$

(b) The argument is identical to the argument under (a), except for the order of the identity matrices. Thus, we conclude that

$$\operatorname{rk}\begin{pmatrix} I_m - BC & O \\ O & I_n \end{pmatrix} = \operatorname{rk}\begin{pmatrix} I_m & O \\ O & I_n - CB \end{pmatrix}$$

and the result follows.

Exercise 5.51 (Upper bound for the rank of a sum) Let A and B be matrices of the same order. We know from Exercise 4.14 that

$$\operatorname{rk}(A + B) \leq \operatorname{rk}(A) + \operatorname{rk}(B).$$

Provide an alternative proof, using partitioned matrices.

Solution
The argument builds on the two matrices

$$Z_1 := \begin{pmatrix} A & O \\ O & B \end{pmatrix} \quad \text{and} \quad Z_2 := \begin{pmatrix} A + B & B \\ B & B \end{pmatrix}.$$

The matrices Z_1 and Z_2 have the same rank, because

$$\begin{pmatrix} I_m & I_m \\ O & I_m \end{pmatrix} \begin{pmatrix} A & O \\ O & B \end{pmatrix} \begin{pmatrix} I_n & O \\ I_n & I_n \end{pmatrix} = \begin{pmatrix} A+B & B \\ B & B \end{pmatrix}.$$

Clearly, $\operatorname{rk}(Z_1) = \operatorname{rk}(A) + \operatorname{rk}(B)$. Also, since $A + B$ is a submatrix of Z_2 we must have $\operatorname{rk}(Z_2) \geq \operatorname{rk}(A + B)$ (Exercise 4.17). Hence,

$$\operatorname{rk}(A + B) \leq \operatorname{rk}(Z_2) = \operatorname{rk}(Z_1) = \operatorname{rk}(A) + \operatorname{rk}(B).$$

Exercise 5.52 (Rank of a 3-by-3 block matrix) Consider the symmetric matrix Z of Exercise 5.19. Show that

$$\operatorname{rk}(Z) = \operatorname{rk}(D) + \operatorname{rk}(E) + \operatorname{rk}(A - BD^{-1}B' - CE^{-1}C')$$

if D and E are nonsingular.

Solution
Let

$$\tilde{A} := A, \quad \tilde{B} := (B : C), \quad \tilde{C} := (B : C)', \quad \tilde{D} := \begin{pmatrix} D & O \\ O & E \end{pmatrix}.$$

Then, using Exercise 5.48,

$$\begin{aligned} \operatorname{rk}(Z) &= \operatorname{rk}(\tilde{D}) + \operatorname{rk}(\tilde{A} - \tilde{B}\tilde{D}^{-1}\tilde{C}) \\ &= \operatorname{rk}(D) + \operatorname{rk}(E) + \operatorname{rk}(A - BD^{-1}B' - CE^{-1}C'). \end{aligned}$$

Exercise 5.53 (Rank of a bordered matrix) Let

$$Z := \begin{pmatrix} 0 & A \\ \alpha & a' \end{pmatrix}.$$

Show that

$$\operatorname{rk}(Z) = \begin{cases} \operatorname{rk}(A) & (\alpha = 0 \text{ and } a \in \operatorname{col}(A')), \\ \operatorname{rk}(A) + 1 & (\text{otherwise}). \end{cases}$$

Solution
If $\alpha \neq 0$ then $\operatorname{rk}(Z) = \operatorname{rk}(A) + 1$ by Exercise 5.46(a). If $\alpha = 0$ then $\operatorname{rk}(Z) = \operatorname{rk}(A' : a)$. If $a \in \operatorname{col}(A')$ then

$$\operatorname{rk}(A' : a) = \operatorname{rk}(A') = \operatorname{rk}(A).$$

If $a \notin \operatorname{col}(A')$ then

$$\operatorname{rk}(A' : a) = \operatorname{rk}(A') + 1 = \operatorname{rk}(A) + 1.$$

5.5 The sweep operator

Exercise 5.54 (Simple sweep) Consider the 2×2 matrix

$$A := \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

(a) Compute $A^{(1)} := \mathrm{SWP}(A, 1)$ and state the condition(s) under which it is defined.
(b) Compute $A^{(2)} := \mathrm{SWP}(A^{(1)}, 2)$ and state the condition(s) under which it is defined.
(c) Show that $A^{(2)} = -A^{-1}$.

Solution
(a) By definition, we have

$$A^{(1)} = \mathrm{SWP}(A, 1) = \begin{pmatrix} -1/a & b/a \\ c/a & d - bc/a \end{pmatrix},$$

provided $a \neq 0$.
(b) Applying the definition to $A^{(1)}$ gives

$$A^{(2)} = \mathrm{SWP}(A^{(1)}, 2) = \frac{a}{ad - bc} \begin{pmatrix} -\frac{ad-bc}{a^2} & \frac{cb}{a^2} & \frac{b}{a} \\ & \frac{c}{a} & -1 \end{pmatrix} = \frac{-1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

provided $a \neq 0$ and $ad - bc \neq 0$.
(c) We recognize $-A^{(2)}$ as the inverse of A or, if we don't, we can verify that $AA^{(2)} = -I_2$.

Exercise 5.55 (General sweep)
(a) Let A be a 3×3 matrix. Compute $\mathrm{SWP}(A, 2)$ and state the condition(s) under which it is defined.
(b) Let A be an $n \times n$ matrix. For $1 \leq p \leq n$, compute $\mathrm{SWP}(A, p)$ and state the condition(s) under which it is defined.

Solution
(a) Let

$$A := \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Then, applying the definition,

$$\mathrm{SWP}(A, 2) = \begin{pmatrix} a_{11} - a_{12}a_{21}/a_{22} & a_{12}/a_{22} & a_{13} - a_{12}a_{23}/a_{22} \\ a_{21}/a_{22} & -1/a_{22} & a_{23}/a_{22} \\ a_{31} - a_{32}a_{21}/a_{22} & a_{32}/a_{22} & a_{33} - a_{32}a_{23}/a_{22} \end{pmatrix},$$

provided $a_{22} \neq 0$.

(b) More generally, if

$$A := \begin{pmatrix} A_{11} & a_{12} & A_{13} \\ a'_{21} & a_{22} & a'_{23} \\ A_{31} & a_{32} & A_{33} \end{pmatrix},$$

where A_{11} has order $p-1$, a_{22} is a scalar, and A_{33} has order $n-p$, then we obtain in the same way

$$\mathrm{SWP}(A,p) = \begin{pmatrix} A_{11} - a_{12}a'_{21}/a_{22} & a_{12}/a_{22} & A_{13} - a_{12}a'_{23}/a_{22} \\ a'_{21}/a_{22} & -1/a_{22} & a'_{23}/a_{22} \\ A_{31} - a_{32}a'_{21}/a_{22} & a_{32}/a_{22} & A_{33} - a_{32}a'_{23}/a_{22} \end{pmatrix},$$

provided a_{22} is nonzero.

***Exercise 5.56 (The sweeping theorem)** Let A be an $n \times n$ matrix and let $1 \leq p \leq n$. Define $A^{(k)}$ recursively by $A^{(k)} := \mathrm{SWP}(A^{(k-1)}, k)$ for $k = 1, \dots, p$ with starting value $A^{(0)} := A$.

(a) If A is partitioned as

$$A := \begin{pmatrix} P & Q \\ R & S \end{pmatrix},$$

where P is a $p \times p$ matrix, show that

$$A^{(p)} = \begin{pmatrix} -P^{-1} & P^{-1}Q \\ RP^{-1} & S - RP^{-1}Q \end{pmatrix}.$$

(b) Hence show that $A^{(n)} = -A^{-1}$.

Solution

(a) We prove this by induction on p. The result is true for $p = 1$, because $A^{(1)} = \mathrm{SWP}(A, 1)$ and the definition of the sweep operator or Exercise 5.54(a). Next, assume that the result holds for $p-1$, and let A be partitioned as

$$A = \begin{pmatrix} A_{11} & a_{12} & A_{13} \\ a'_{21} & a_{22} & a'_{23} \\ A_{31} & a_{32} & A_{33} \end{pmatrix},$$

where A_{11} has order $p-1$, a_{22} is a scalar, and A_{33} has order $n-p$. Then, by the induction hypothesis, we have

$$A^{(p-1)} = \begin{pmatrix} -A_{11}^{-1} & A_{11}^{-1}a_{12} & A_{11}^{-1}A_{13} \\ a'_{21}A_{11}^{-1} & a_{22} - a'_{21}A_{11}^{-1}a_{12} & a'_{23} - a'_{21}A_{11}^{-1}A_{13} \\ A_{31}A_{11}^{-1} & a_{32} - A_{31}A_{11}^{-1}a_{12} & A_{33} - A_{31}A_{11}^{-1}A_{13} \end{pmatrix}.$$

We now use Exercise 5.55(b); this shows that $\mathrm{SWP}(A^{(p-1)}, p)$ is equal to

$$\begin{pmatrix} -B_{11} & -b_{12} & B_{11}A_{13} + b_{12}a'_{23} \\ -b'_{21} & -b_{22} & b'_{21}A_{13} + b_{22}a'_{23} \\ A_{31}B_{11} + a_{32}b'_{21} & A_{31}b_{12} + a_{32}b_{22} & A_{33} - D \end{pmatrix},$$

where

$$B_{11} := A_{11}^{-1} + A_{11}^{-1}a_{12}a_{21}'A_{11}^{-1}/\beta,$$
$$b_{12} := -A_{11}^{-1}a_{12}/\beta, \quad b_{21}' := -a_{21}'A_{11}^{-1}/\beta,$$
$$b_{22} := 1/\beta, \quad \beta := a_{22} - a_{21}'A_{11}^{-1}a_{12},$$
$$D := A_{31}A_{11}^{-1}A_{13} + (a_{32} - A_{31}A_{11}^{-1}a_{12})(a_{23}' - a_{21}'A_{11}^{-1}A_{13})/\beta.$$

Noticing that

$$\begin{pmatrix} A_{11} & a_{12} \\ a_{21}' & a_{22} \end{pmatrix}^{-1} = \begin{pmatrix} B_{11} & b_{12} \\ b_{21}' & b_{22} \end{pmatrix},$$

using Exercise 5.16(a), and that

$$D = (A_{31} : a_{32}) \begin{pmatrix} B_{11} & b_{12} \\ b_{21}' & b_{22} \end{pmatrix} \begin{pmatrix} A_{13} \\ a_{23}' \end{pmatrix},$$

the result follows.
(b) This follows directly from (a). The inverse of A can thus be computed by n sequential sweep operations, a very useful fact in numerical inversion routines.

Exercise 5.57 (Sweeping and linear equations)
(a) Show how the sweep operator can be used to solve the linear system $PX = Q$ for nonsingular P.
(b) In particular, solve the system $2x_1 + 3x_2 = 8$ and $4x_1 + 5x_2 = 14$ using the sweep operator.

Solution
(a) We know from Exercise 5.56 that

$$A := \begin{pmatrix} P & Q \\ R & S \end{pmatrix} \implies A^{(p)} = \begin{pmatrix} -P^{-1} & P^{-1}Q \\ RP^{-1} & S - RP^{-1}Q \end{pmatrix}.$$

Hence, the solution $P^{-1}Q$ appears as the $(1,2)$-block of $A^{(p)}$, where p denotes the order of the square matrix P.
(b) Denoting irrelevant elements by $*$s, we define

$$A^{(0)} := \begin{pmatrix} 2 & 3 & 8 \\ 4 & 5 & 14 \\ * & * & * \end{pmatrix}.$$

This gives

$$A^{(1)} := \mathrm{SWP}(A^{(0)}, 1) = \begin{pmatrix} -1/2 & 3/2 & 4 \\ 2 & -1 & -2 \\ * & * & * \end{pmatrix}$$

and

$$\boldsymbol{A}^{(2)} := \mathrm{SWP}(\boldsymbol{A}^{(1)}, 2) = \begin{pmatrix} 5/2 & -3/2 & 1 \\ -2 & 1 & 2 \\ * & * & * \end{pmatrix},$$

so that the solution is $x_1 = 1$, $x_2 = 2$.

Notes

A good survey of results with partitioned matrices can be found in Chapter 2 of Zhang (1999). The inequalities in Exercise 5.47 were first obtained by Sylvester in 1884 and Frobenius in 1911. Sylvester's inequality is called the "law of nullity", because it implies that

$$\dim(\ker(\boldsymbol{AB})) \leq \dim(\ker(\boldsymbol{A})) + \dim(\ker(\boldsymbol{B})),$$

and the dimension of the kernel of a matrix is known as its "nullity". The sweep operator (Exercises 5.54–5.57) plays a role in inversion routines. It was introduced by Beaton (1964); see also Dempster (1969).