

## 4.3 Rank, Equivalence, and Factorizations

Similarity compares two matrices of one operator, where the same basis is used for input and output, and the last section showed that deciding it is hard. This section asks what happens when the two bases may be chosen independently, as they may for any map  $T: V \rightarrow W$ . The answer is strikingly simple: the rank is then the only thing that matters. On the way we meet the column–row factorization, which writes every matrix as a product through a space of dimension equal to its rank, and we collect the basic inequalities for the rank of a composition.

### 4.3.1 The column–row factorization

In Chapter 3 we computed the rank of a matrix by row reduction, and the reduction gave more than a number. It produced a basis of the column space, the pivot columns of  $\mathbf{A}$  (Theorem 3.5.2), and it expressed every other column in terms of them, with coefficients read off the reduced row echelon form. Recording both pieces of information as matrices gives a factorization.

#### Theorem 4.3.1. Column–Row Factorization

Let  $\mathbf{A} \in M_{m \times n}(F)$  have rank  $r \geq 1$ , with pivot columns  $j_1 < \cdots < j_r$ . Let  $\mathbf{C} \in M_{m \times r}(F)$  be the matrix whose columns are the pivot columns  $\mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_r}$  of  $\mathbf{A}$ , and let  $\mathbf{R} \in M_{r \times n}(F)$  consist of the first  $r$  rows of the reduced row echelon form of  $\mathbf{A}$ , that is, its nonzero rows. Then

$$\mathbf{A} = \mathbf{C}\mathbf{R}.$$

Moreover, the columns of  $\mathbf{C}$  are linearly independent and the rows of  $\mathbf{R}$  are linearly independent.

*Idea.* Column  $l$  of  $\mathbf{C}\mathbf{R}$  is  $\mathbf{C}$  times column  $l$  of  $\mathbf{R}$ , that is, a combination of the pivot columns of  $\mathbf{A}$  with the entries of column  $l$  of the RREF as coefficients. For a pivot column the coefficients are a standard basis vector, and for a free column they are exactly the coefficients that Theorem 3.5.2 (b) provides. So the factorization is that theorem written as a matrix identity.

*Proof.* Let  $\mathbf{E}$  be the RREF of  $\mathbf{A}$ , with entries  $e_{kl}$ . By Theorem 3.5.5,  $\mathbf{E}$  has exactly  $r$  pivots, so its nonzero rows are rows  $1, \dots, r$ . Fix a column index  $l$ . By Theorem 0.9.3, column  $l$  of  $\mathbf{C}\mathbf{R}$  is

$$\mathbf{C} \begin{pmatrix} e_{1l} \\ \vdots \\ e_{rl} \end{pmatrix} = e_{1l}\mathbf{a}_{j_1} + \cdots + e_{rl}\mathbf{a}_{j_r}.$$

*Case 1:*  $l = j_i$  is a pivot column. By Lemma 3.3.3 (a), column  $j_i$  of  $\mathbf{E}$  is  $\mathbf{e}_i$ , so the combination is  $\mathbf{a}_{j_i} = \mathbf{a}_l$ .

*Case 2:*  $l$  is a free column. Let  $t$  be the number of pivot columns to the left of  $l$ . For  $k > t$ , the leading entry of row  $k$  of  $\mathbf{E}$  lies in column  $j_k > l$ , so  $e_{kl} = 0$ . The combination is therefore  $e_{1l}\mathbf{a}_{j_1} + \cdots + e_{tl}\mathbf{a}_{j_t}$ , which equals  $\mathbf{a}_l$  by Theorem 3.5.2 (b).

In both cases column  $l$  of  $\mathbf{C}\mathbf{R}$  is column  $l$  of  $\mathbf{A}$ ; hence  $\mathbf{A} = \mathbf{C}\mathbf{R}$ . The columns of  $\mathbf{C}$  form a basis of  $\text{col}(\mathbf{A})$  by Theorem 3.5.2 (a), so they are independent. The rows of  $\mathbf{R}$  are the

nonzero rows of  $\mathbf{E}$ , which form a basis of  $\text{row}(\mathbf{A})$  by Theorem 3.5.3, so they are independent. This proves the theorem. ■

In words: every matrix of rank  $r$  is a “tall” matrix with independent columns times a “wide” matrix with independent rows, and the middle size is exactly  $r$ . The factorization cannot be squeezed further.

### Proposition 4.3.2. Rank as the size of the thinnest factorization

Let  $\mathbf{A} \in M_{m \times n}(F)$  be nonzero. Then  $\text{rank } \mathbf{A}$  is the **smallest** integer  $s \geq 1$  such that  $\mathbf{A} = \mathbf{B}\mathbf{D}$  for some  $\mathbf{B} \in M_{m \times s}(F)$  and  $\mathbf{D} \in M_{s \times n}(F)$ .

*Proof.* Let  $r = \text{rank } \mathbf{A}$ ; then  $r \geq 1$  since  $\mathbf{A} \neq \mathbf{O}$ . By Theorem 4.3.1, a factorization with  $s = r$  exists. Conversely, suppose  $\mathbf{A} = \mathbf{B}\mathbf{D}$  with  $\mathbf{B} \in M_{m \times s}(F)$ . By Theorem 3.5.9,  $r = \text{rank}(\mathbf{B}\mathbf{D}) \leq \text{rank } \mathbf{B}$ , and  $\text{rank } \mathbf{B} \leq s$  because  $\text{col}(\mathbf{B})$  is spanned by  $s$  columns (Theorem 1.6.3). Hence  $r \leq s$ . ■

This gives a second view of “row rank equals column rank”. Let  $\mathbf{A} \neq \mathbf{O}$ , and let  $\mathbf{B}$  be the matrix whose  $s = \dim \text{col}(\mathbf{A})$  columns form a basis of  $\text{col}(\mathbf{A})$ . Each column  $\mathbf{a}_j$  of  $\mathbf{A}$  is  $\mathbf{B}\mathbf{d}_j$  for its coordinate column  $\mathbf{d}_j$  (Theorem 0.9.3), so  $\mathbf{A} = \mathbf{B}\mathbf{D}$  with  $\mathbf{D} = (\mathbf{d}_1 \ \cdots \ \mathbf{d}_n)$ . Then every row of  $\mathbf{A}$  is a row of  $\mathbf{B}$  times  $\mathbf{D}$ , a combination of the  $s$  rows of  $\mathbf{D}$  (Theorem 0.9.2 (3) and the row form of Theorem 0.9.3), so  $\dim \text{row}(\mathbf{A}) \leq s = \dim \text{col}(\mathbf{A})$ . Applying the same argument to  $\mathbf{A}^\top$  gives the reverse inequality. In the language of maps, the factorization says that  $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$  can be done in two steps,  $F^n \rightarrow F^r \rightarrow F^m$ : first a surjection onto a space of dimension  $r$ , then an injection.

### Example. A column–row factorization

Find the column–row factorization of  $\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 2 & 4 & 1 & 4 \\ 3 & 6 & 1 & 7 \end{pmatrix} \in M_{3 \times 4}(\mathbb{R})$ .

*Solution.* Row reduce. Subtracting 2 times row 1 from row 2 and 3 times row 1 from row 3 gives rows  $(1, 2, 0, 3)$ ,  $(0, 0, 1, -2)$ ,  $(0, 0, 1, -2)$ . Subtracting row 2 from row 3 gives the RREF

$$\begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The pivot columns are columns 1 and 3, so  $r = 2$ , and by Theorem 4.3.1

$$\mathbf{A} = \mathbf{C}\mathbf{R} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{pmatrix}.$$

Check by rows: row 2 of  $\mathbf{CR}$  is  $2(1, 2, 0, 3) + (0, 0, 1, -2) = (2, 4, 1, 4)$ , and row 3 is  $3(1, 2, 0, 3) + (0, 0, 1, -2) = (3, 6, 1, 7)$ . The last column of  $\mathbf{R}$  says  $\mathbf{a}_4 = 3\mathbf{a}_1 - 2\mathbf{a}_3$ , and indeed  $3(1, 2, 3) - 2(0, 1, 1) = (3, 4, 7)$ .

## 4.3.2 Equivalent matrices

Go back to the change-of-basis square for a map  $T: V \rightarrow W$  (Theorem 4.2.3):

$$[T]_{\mathcal{B}'}^{\mathcal{C}'} = [\text{id}]_{\mathcal{C}}^{\mathcal{C}'} [T]_{\mathcal{B}}^{\mathcal{C}} [\text{id}]_{\mathcal{B}'}^{\mathcal{B}}.$$

When  $V \neq W$ , nothing ties the basis of  $V$  to the basis of  $W$ . So the two outer matrices are **unrelated** invertible matrices, of sizes  $m \times m$  and  $n \times n$ . Any two matrices of the same map are related in this way, and this recurring shape deserves a name.

*Two matrices are equivalent when they describe the same linear map, with the input basis and the output basis each changed freely.*

### Definition 4.3.1. Equivalent matrices

Let  $\mathbf{A}, \mathbf{B} \in M_{m \times n}(F)$ . We say  $\mathbf{A}$  is **equivalent** to  $\mathbf{B}$  if there **exist** invertible matrices  $\mathbf{Q} \in M_m(F)$  and  $\mathbf{P} \in M_n(F)$  such that

$$\mathbf{B} = \mathbf{QAP}.$$

In words: we may multiply on the left by **any** invertible matrix and on the right by **any** invertible matrix, and the two need not be related. Both sizes are forced:  $\mathbf{Q}$  acts on the  $m$  rows and  $\mathbf{P}$  on the  $n$  columns. This is an equivalence relation on  $M_{m \times n}(F)$ :  $\mathbf{A} = \mathbf{I}_m \mathbf{A} \mathbf{I}_n$ ; if  $\mathbf{B} = \mathbf{QAP}$  then  $\mathbf{A} = \mathbf{Q}^{-1} \mathbf{B} \mathbf{P}^{-1}$ ; and if also  $\mathbf{C} = \mathbf{Q}' \mathbf{B} \mathbf{P}'$ , then  $\mathbf{C} = (\mathbf{Q}' \mathbf{Q}) \mathbf{A} (\mathbf{P} \mathbf{P}')$ , where products of invertible matrices are invertible by Theorem 0.9.7 (part 3).

### Examples.

- **Row equivalence.** If  $\mathbf{B}$  is obtained from  $\mathbf{A}$  by row operations, then  $\mathbf{B} = \mathbf{EA}$  with  $\mathbf{E}$  invertible (Theorem 3.4.8), so  $\mathbf{A}$  and  $\mathbf{B}$  are equivalent with  $\mathbf{Q} = \mathbf{E}$  and  $\mathbf{P} = \mathbf{I}_n$ . Column operations correspond to multiplying on the right, and they are allowed too.
- **Similarity.** If  $\mathbf{B} = \mathbf{P}^{-1} \mathbf{A} \mathbf{P}$ , then  $\mathbf{A}$  and  $\mathbf{B}$  are equivalent, with  $\mathbf{Q} = \mathbf{P}^{-1}$ .
- **The identity and a swap.**  $\mathbf{I}_2$  and  $S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  are equivalent, since  $S = S \mathbf{I}_2 \mathbf{I}_2$ . They are not similar, because  $\mathbf{I}_2$  is similar only to itself.
- **Degenerate case.** The zero matrix  $\mathbf{O}_{m \times n}$  is equivalent only to itself, since  $\mathbf{QOP} = \mathbf{O}$ . It is the matrix of the zero map in every pair of bases.

**Nonexample by minimal change.** Keep  $\mathbf{B} = \mathbf{QAP}$  but drop the requirement that  $\mathbf{Q}$  be invertible. Then every matrix could be sent to the zero matrix with  $\mathbf{Q} = \mathbf{O}$ , and the relation would not even be symmetric:  $\mathbf{O} = \mathbf{O} \cdot \mathbf{I}_2 \cdot \mathbf{I}_2$ , but  $\mathbf{I}_2 \neq \mathbf{QOP}$  for any  $\mathbf{Q}, \mathbf{P}$ . The clause “invertible” is what makes  $\mathbf{QAP}$  a change of basis rather than an arbitrary composition.

*Warning. Equivalent is much weaker than similar.* The matrices  $\mathbf{I}_2$  and  $J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  are equivalent, since  $J = J \mathbf{I}_2 \mathbf{I}_2$  and  $J$  is invertible, but they are not similar (see the warning

in the previous section). For square matrices, “equivalent” allows a different change of basis on each side, and that freedom destroys almost all the structure an operator has.

How much structure survives? The rank does, since multiplying by invertible matrices does not change it (Theorem 3.5.9). The theorem below says that nothing else survives. The target is the simplest matrix of each rank.

For  $0 \leq r \leq \min(m, n)$ , let  $\mathbf{N}_r \in M_{m \times n}(F)$  be the block matrix

$$\mathbf{N}_r := \begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix},$$

with 1 in positions  $(1, 1), \dots, (r, r)$  and 0 everywhere else. Its first  $r$  columns are  $\mathbf{e}_1, \dots, \mathbf{e}_r \in F^m$ , and its remaining columns are zero.

### Theorem 4.3.3. Rank Normal Form

- (a) Let  $V$  and  $W$  be finite-dimensional with  $\dim V = n$  and  $\dim W = m$ , and let  $T: V \rightarrow W$  be linear of rank  $r$ . Then there are bases  $\mathcal{B}$  of  $V$  and  $\mathcal{C}$  of  $W$  such that  $[T]_{\mathcal{B}}^{\mathcal{C}} = \mathbf{N}_r$ .
- (b) Every  $\mathbf{A} \in M_{m \times n}(F)$  is equivalent to  $\mathbf{N}_r$ , where  $r = \text{rank } \mathbf{A}$ .
- (c) Two matrices  $\mathbf{A}, \mathbf{B} \in M_{m \times n}(F)$  are equivalent if and only if  $\text{rank } \mathbf{A} = \text{rank } \mathbf{B}$ .

*Idea.* Work backwards from the goal.  $[T]_{\mathcal{B}}^{\mathcal{C}} = \mathbf{N}_r$  says:  $T$  sends the first  $r$  vectors of  $\mathcal{B}$  to the first  $r$  vectors of  $\mathcal{C}$ , and sends the last  $n - r$  vectors of  $\mathcal{B}$  to 0. So ① the last  $n - r$  vectors of  $\mathcal{B}$  should be a basis of  $\ker T$ , and by Rank–Nullity that is exactly the right number; ② the first  $r$  vectors of  $\mathcal{B}$  extend it to a basis of  $V$ ; ③ their images are forced to be the first  $r$  vectors of  $\mathcal{C}$ , and we must show they are independent, which follows by counting; ④ extend them to a basis of  $W$ . These are the bases of the proof of Rank–Nullity, one on each side.

*Proof.*

- (a) By the Rank–Nullity Theorem (Theorem 2.3.1),  $\dim \ker T = n - r$ . Let  $(u_{r+1}, \dots, u_n)$  be a basis of  $\ker T$ , a subspace of  $V$  by Theorem 2.2.1. By the Basis Extension Theorem (Theorem 1.6.5) there are  $u_1, \dots, u_r \in V$  such that  $\mathcal{B} = (u_1, \dots, u_r, u_{r+1}, \dots, u_n)$  is a basis of  $V$ ; the new vectors may be listed first, since reordering a basis gives a basis. Put  $w_i = Tu_i$  for  $i = 1, \dots, r$ .  
By Theorem 2.2.4,  $\text{im } T = \text{span}(Tu_1, \dots, Tu_n)$ . Since  $Tu_j = 0$  for  $j > r$ , removing these vectors does not change the span, so  $(w_1, \dots, w_r)$  spans  $\text{im } T$ . This list has length  $r = \dim \text{im } T$ , so it is a basis of  $\text{im } T$  by Theorem 1.6.4 (b); in particular it is linearly independent in  $W$ . By the Basis Extension Theorem there are  $w_{r+1}, \dots, w_m \in W$  such that  $\mathcal{C} = (w_1, \dots, w_m)$  is a basis of  $W$ .  
Now read off the columns of  $[T]_{\mathcal{B}}^{\mathcal{C}}$  (Definition 4.1.1). For  $i \leq r$ ,  $Tu_i = w_i$ , whose  $\mathcal{C}$ -coordinate vector is  $\mathbf{e}_i$ . For  $j > r$ ,  $Tu_j = 0$ , whose coordinate vector is  $\mathbf{0}$ . These are the columns of  $\mathbf{N}_r$ .
- (b) Let  $T: F^n \rightarrow F^m$ ,  $T\mathbf{x} = \mathbf{A}\mathbf{x}$ , and let  $\mathcal{E}_n, \mathcal{E}_m$  be the standard bases. The  $j$ -th column of  $[T]_{\mathcal{E}_n}^{\mathcal{E}_m}$  is  $[\mathbf{A}\mathbf{e}_j]_{\mathcal{E}_m} = \mathbf{A}\mathbf{e}_j$ , the  $j$ -th column of  $\mathbf{A}$  (Theorem 0.9.3), so  $[T]_{\mathcal{E}_n}^{\mathcal{E}_m} = \mathbf{A}$ . By Theorem 4.1.5,  $\text{rank } T = \text{rank } \mathbf{A} = r$ . Take  $\mathcal{B}, \mathcal{C}$  as in (a). By Theorem 4.2.3,

$$\mathbf{N}_r = [T]_{\mathcal{B}}^{\mathcal{C}} = [\text{id}]_{\mathcal{E}_m}^{\mathcal{C}} \mathbf{A} [\text{id}]_{\mathcal{B}}^{\mathcal{E}_n},$$

and both outer matrices are invertible by Theorem 4.2.1 (c). Hence  $\mathbf{A}$  is equivalent to  $\mathbf{N}_r$ .

- (c) ( $\Rightarrow$ ) If  $\mathbf{B} = \mathbf{QAP}$  with  $\mathbf{Q}, \mathbf{P}$  invertible, then  $\text{rank } \mathbf{B} = \text{rank } \mathbf{A}$  by Theorem 3.5.9.  
 ( $\Leftarrow$ ) If  $\text{rank } \mathbf{A} = \text{rank } \mathbf{B} = r$ , then by (b) both  $\mathbf{A}$  and  $\mathbf{B}$  are equivalent to  $\mathbf{N}_r$ . Since equivalence is symmetric and transitive,  $\mathbf{A}$  is equivalent to  $\mathbf{B}$ . This proves the theorem. ■

So the equivalence classes of  $M_{m \times n}(F)$  are exactly the sets of matrices of rank  $r$ , for  $r = 0, 1, \dots, \min(m, n)$ : there are  $\min(m, n) + 1$  of them, whatever the field. Every linear map between finite-dimensional spaces looks like “keep  $r$  coordinates and delete the rest” once the bases on both sides are chosen well. **The right pair of bases makes any single map trivial.**

*Remark.* Contrast this with operators. For  $T \in \mathcal{L}(V)$  we must use **one** basis for input and output, and the relation becomes similarity. There the rank is far from the whole story:  $\mathbf{I}_2$  and  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  have the same rank but are not similar. Classifying operators up to a change of one basis is a much harder problem, answered in Chapter 10 by the Jordan and rational canonical forms.

The proof is also an algorithm. To find  $\mathbf{Q}$  and  $\mathbf{P}$  for a concrete matrix, compute a basis of  $\text{nul}(\mathbf{A})$ , extend it, and take images.

**Example. Bringing a  $2 \times 3$  matrix to normal form**

Let  $\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \end{pmatrix} \in M_{2 \times 3}(\mathbb{R})$ . Find invertible  $\mathbf{Q} \in M_2(\mathbb{R})$  and  $\mathbf{P} \in M_3(\mathbb{R})$  with  $\mathbf{QAP} = \mathbf{N}_r$ .

*Solution.* Subtracting 2 times row 1 from row 2 gives  $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ , so  $r = \text{rank } \mathbf{A} = 2$ , and the solutions of  $\mathbf{Ax} = \mathbf{0}$  satisfy  $x_3 = 0$ ,  $x_1 = -2x_2$ . Hence  $\text{nul}(\mathbf{A}) = \text{span}((-2, 1, 0))$ , as Rank–Nullity predicts ( $3 - 2 = 1$ ).

Following the proof, let  $\mathbf{u}_3 = (-2, 1, 0)$ , and extend with  $\mathbf{u}_1 = \mathbf{e}_1$ ,  $\mathbf{u}_2 = \mathbf{e}_3$ . The list  $(\mathbf{e}_1, \mathbf{e}_3, \mathbf{u}_3)$  is independent: in  $a\mathbf{e}_1 + b\mathbf{e}_3 + c(-2, 1, 0) = (a - 2c, c, b) = \mathbf{0}$ , the second and third entries give  $c = b = 0$ , and then  $a = 0$ . So it is a basis  $\mathcal{B}$  of  $\mathbb{R}^3$ . The images are  $\mathbf{w}_1 = \mathbf{A}\mathbf{e}_1 = (1, 2)$  and  $\mathbf{w}_2 = \mathbf{A}\mathbf{e}_3 = (1, 3)$ , which already form a basis  $\mathcal{C}$  of  $\mathbb{R}^2$  since  $r = m = 2$ . Therefore

$$\mathbf{P} = [\text{id}]_{\mathcal{B}}^{\mathcal{E}_3} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{Q} = [\text{id}]_{\mathcal{E}_2}^{\mathcal{C}} = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix},$$

using Theorem 4.2.1 (c) and Theorem 0.9.8. Check:  $\mathbf{AP} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{pmatrix}$  (the columns are  $\mathbf{w}_1$ ,

$\mathbf{w}_2, \mathbf{0})$ , and

$$\mathbf{QAP} = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \mathbf{N}_2.$$

The pair  $(\mathbf{Q}, \mathbf{P})$  is far from unique; any choice of extension gives another.

*Quick check.* How many equivalence classes does  $M_{3 \times 2}(F)$  have, and which is the class of  $\begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \end{pmatrix}$ ?

*Solution.* Ranks in  $M_{3 \times 2}(F)$  range over 0, 1, 2, so there are  $\min(3, 2) + 1 = 3$  classes by Theorem 4.3.3 (c). The given matrix is nonzero and its second column is twice the first, so its rank is 1, and it lies in the class of  $\mathbf{N}_1$ : all matrices of rank 1.

### 4.3.3 Rank inequalities for compositions

The rank of a composition cannot exceed the rank of either factor, and it cannot fall too far either. Both facts come from the same observation:  $ST$  is  $S$  applied to the subspace  $\text{im } T$ . For a subspace  $X \subseteq W$  and a linear map  $S: W \rightarrow U$ , the **restriction**  $S|_X: X \rightarrow U$ ,  $x \mapsto Sx$ , is linear, since its values are those of  $S$ . Its kernel is  $\ker S \cap X$ .

#### Theorem 4.3.4. Rank of a Composition

Let  $T: V \rightarrow W$  and  $S: W \rightarrow U$  be linear, with  $V$  and  $W$  finite-dimensional. Then

$$\text{rank}(ST) \leq \min(\text{rank } S, \text{rank } T).$$

*Proof.* Consider the restriction  $S' = S|_{\text{im } T}: \text{im } T \rightarrow U$ . For  $v \in V$ ,  $(ST)(v) = S(Tv) = S'(Tv)$ , so  $\text{im}(ST) \subseteq \text{im } S'$ ; conversely every  $S'w$  with  $w = Tv \in \text{im } T$  equals  $(ST)(v)$ . Hence  $\text{im}(ST) = \text{im } S'$ . Since  $\text{im } T$  is finite-dimensional (a subspace of  $W$ , Theorem 1.6.7), Rank–Nullity (Theorem 2.3.1) applied to  $S'$  gives

$$\text{rank}(ST) = \dim \text{im } S' = \dim \text{im } T - \dim \ker S' \leq \text{rank } T.$$

Also  $\text{im}(ST) = \text{im } S' \subseteq \text{im } S$ , and  $\text{im } S$  is finite-dimensional by Rank–Nullity for  $S$ , so  $\text{rank}(ST) \leq \text{rank } S$  by Theorem 1.6.7. This proves the theorem. ■

The lower bound uses the same restriction, but now we control its kernel instead of ignoring it.

#### Theorem 4.3.5. Sylvester’s Rank Inequality

Let  $T: V \rightarrow W$  and  $S: W \rightarrow U$  be linear, with  $V$  and  $W$  finite-dimensional. Then

$$\text{rank}(ST) \geq \text{rank } S + \text{rank } T - \dim W.$$

In matrix form: for  $\mathbf{A} \in M_{m \times n}(F)$  and  $\mathbf{B} \in M_{n \times p}(F)$ ,  $\text{rank}(\mathbf{AB}) \geq \text{rank } \mathbf{A} + \text{rank } \mathbf{B} - n$ .

*Idea.* Of the  $\text{rank } T$  dimensions that reach  $W$ , the map  $S$  can destroy only those lying in  $\ker S$ , and  $\ker S$  has dimension  $\dim W - \text{rank } S$ . Rank–Nullity for the restriction  $S|_{\text{im } T}$  turns this sentence into an inequality.

*Proof.* Let  $S' = S|_{\text{im } T}$ , so that  $\text{im } S' = \text{im}(ST)$  as in the proof of Theorem 4.3.4, and  $\ker S' = \ker S \cap \text{im } T \subseteq \ker S$ . By Rank–Nullity for  $S'$  and then for  $S$ ,

$$\begin{aligned} \text{rank } T &= \text{rank}(ST) + \dim(\ker S \cap \text{im } T) \\ &\leq \text{rank}(ST) + \dim \ker S = \text{rank}(ST) + \dim W - \text{rank } S, \end{aligned}$$

where the inequality is Theorem 1.6.7. Rearranging gives the claim. For matrices, apply it to  $S\mathbf{y} = \mathbf{A}\mathbf{y}$  on  $W = F^n$  and  $T\mathbf{x} = \mathbf{B}\mathbf{x}$  on  $V = F^p$ : the matrix of  $ST$  in standard bases is  $\mathbf{AB}$  by Theorem 4.1.2, and ranks of maps and matrices agree by Theorem 4.1.5. ■

If you worked Exercise (C1: A Lower Bound for the Rank of a Composite), you have seen this proof already; we record the result here as a theorem so that later arguments can cite it. A typical use: if  $\mathbf{AB} = \mathbf{O}$  with  $\mathbf{A} \in M_{m \times n}(F)$  and  $\mathbf{B} \in M_{n \times p}(F)$ , then  $0 \geq \text{rank } \mathbf{A} + \text{rank } \mathbf{B} - n$ , that is,

$$\text{rank } \mathbf{A} + \text{rank } \mathbf{B} \leq n.$$

For instance, there are no  $\mathbf{A}, \mathbf{B} \in M_3(\mathbb{R})$  of rank 2 with  $\mathbf{AB} = \mathbf{O}$ .

**Warning.** The rank of a product depends on the order of the factors. For  $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  and  $\mathbf{B} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ,  $\mathbf{AB} = \mathbf{B}$  has rank 1 but  $\mathbf{BA} = \mathbf{O}$  has rank 0. Theorem 3.5.9 and Sylvester’s inequality bound  $\text{rank}(\mathbf{AB})$  from both sides; they do not determine it, and they say nothing that would make  $\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{BA})$ .

## 4.3.4 Exercises

### 4.3.4.1 A. Check your understanding

#### Exercise. A1

- Define what it means for  $\mathbf{A}, \mathbf{B} \in M_{m \times n}(F)$  to be equivalent.
- State the Rank Normal Form theorem for a matrix  $\mathbf{A} \in M_{m \times n}(F)$ .
- True or false: equivalent square matrices are similar. Justify your answer.
- True or false:  $\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{BA})$  for all  $\mathbf{A}, \mathbf{B} \in M_n(F)$ . Justify your answer.

- (e) How many equivalence classes does  $M_{2 \times 5}(\mathbb{F}_7)$  have?
- (f) In the proof of the normal form for  $T: V \rightarrow W$ , which vectors form the end of the basis of  $V$ , and which vectors form the start of the basis of  $W$ ?

*Solution.*

- (a) There are invertible  $\mathbf{Q} \in M_m(F)$  and  $\mathbf{P} \in M_n(F)$  with  $\mathbf{B} = \mathbf{QAP}$  (Definition 4.3.1).
- (b) If  $r = \text{rank } \mathbf{A}$ , there are invertible  $\mathbf{Q}, \mathbf{P}$  with  $\mathbf{QAP} = \begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$ ; and two  $m \times n$  matrices are equivalent if and only if they have the same rank (Theorem 4.3.3).
- (c) False.  $\mathbf{I}_2$  and  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  both have rank 2, so they are equivalent, but  $\mathbf{I}_2$  is similar only to itself.
- (d) False. With  $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  and  $\mathbf{B} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ,  $\text{rank}(\mathbf{AB}) = 1$  and  $\text{rank}(\mathbf{BA}) = 0$ .
- (e) The possible ranks are 0, 1, 2, so there are 3 classes. The field plays no role.
- (f) The last  $n - r$  vectors of the basis of  $V$  form a basis of  $\ker T$ . The first  $r$  vectors of the basis of  $W$  are the images of the first  $r$  basis vectors of  $V$ , and they form a basis of  $\text{im } T$ .

#### 4.3.4.2 B. Practice

##### Exercise. B1: A column–row factorization

Find the column–row factorization of

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 & 0 \\ 2 & -2 & 5 & 1 \\ 1 & -1 & 3 & 1 \end{pmatrix} \in M_{3 \times 4}(\mathbb{R}).$$

Hence write the second and fourth columns of  $\mathbf{A}$  as combinations of its pivot columns.

*Solution.* Subtract 2 times row 1 from row 2 and row 1 from row 3: the rows become  $(1, -1, 2, 0)$ ,  $(0, 0, 1, 1)$ ,  $(0, 0, 1, 1)$ . Subtract row 2 from row 3, and then 2 times row 2 from row 1:

$$\text{RREF} = \begin{pmatrix} 1 & -1 & 0 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The pivot columns are 1 and 3, so  $\text{rank } \mathbf{A} = 2$ , and by Theorem 4.3.1

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & -2 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Check by rows:  $2(1, -1, 0, -2) + 5(0, 0, 1, 1) = (2, -2, 5, 1)$  and  $(1, -1, 0, -2) + 3(0, 0, 1, 1) = (1, -1, 3, 1)$ ; row 1 is  $(1, -1, 0, -2) + 2(0, 0, 1, 1)$ . Reading columns 2 and 4 of  $\mathbf{R}$ :  $\mathbf{a}_2 = -\mathbf{a}_1$  and  $\mathbf{a}_4 = -2\mathbf{a}_1 + \mathbf{a}_3$ . Indeed  $-2(1, 2, 1) + (2, 5, 3) = (0, 1, 1) = \mathbf{a}_4$ .

**Exercise. B2: Normal form of a rank-one matrix**

Let  $\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \end{pmatrix} \in M_{2 \times 3}(\mathbb{R})$ . Find invertible  $\mathbf{Q} \in M_2(\mathbb{R})$  and  $\mathbf{P} \in M_3(\mathbb{R})$  such that  $\mathbf{QAP} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ , following the proof of [Theorem 4.3.3](#).

*Solution.* Row 2 is  $-2$  times row 1, and  $\mathbf{A} \neq \mathbf{0}$ , so  $\text{rank } \mathbf{A} = 1$ . The null space is  $\{x_1 - x_2 + 2x_3 = 0\}$ ; setting the free variables  $x_2, x_3$  to  $(1, 0)$  and  $(0, 1)$  gives the basis  $\mathbf{u}_2 = (1, 1, 0)$ ,  $\mathbf{u}_3 = (-2, 0, 1)$ . Extend with  $\mathbf{u}_1 = \mathbf{e}_1$ : in  $a\mathbf{e}_1 + b\mathbf{u}_2 + c\mathbf{u}_3 = (a + b - 2c, b, c) = \mathbf{0}$ , the last two entries give  $b = c = 0$  and then  $a = 0$ , so  $\mathcal{B} = (\mathbf{e}_1, \mathbf{u}_2, \mathbf{u}_3)$  is a basis of  $\mathbb{R}^3$ . Its first image is  $\mathbf{w}_1 = \mathbf{A}\mathbf{e}_1 = (1, -2)$ , which we extend to the basis  $\mathcal{C} = ((1, -2), \mathbf{e}_2)$  of  $\mathbb{R}^2$  (independent, since  $a(1, -2) + b(0, 1) = (a, -2a + b) = \mathbf{0}$  forces  $a = b = 0$ ). Then

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

Check:  $\mathbf{AP}$  has columns  $\mathbf{A}\mathbf{e}_1 = (1, -2)$ ,  $\mathbf{A}\mathbf{u}_2 = \mathbf{0}$ ,  $\mathbf{A}\mathbf{u}_3 = \mathbf{0}$ , and  $\mathbf{Q}(1, -2) = (1, 0)$ . So  $\mathbf{QAP} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ . Both  $\mathbf{Q}$  and  $\mathbf{P}$  are invertible by Theorem 4.2.1 (c).

**Exercise. B3: Rank-one pieces**

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

- Use [Theorem 4.3.1](#) to show that if  $\text{rank } \mathbf{A} = 1$ , then  $\mathbf{A} = \mathbf{xy}^\top$  for some nonzero  $\mathbf{x} \in F^m$  and  $\mathbf{y} \in F^n$ .
- Show that if  $\text{rank } \mathbf{A} = r \geq 1$ , then  $\mathbf{A}$  is a sum of  $r$  matrices of rank 1, and that it is not a sum of fewer than  $r$  matrices of rank 1.
- Write  $\begin{pmatrix} 1 & 2 & 0 & 3 \\ 2 & 4 & 1 & 4 \\ 3 & 6 & 1 & 7 \end{pmatrix}$  as a sum of two rank-one matrices.

*Solution.*

- With  $r = 1$ ,  $\mathbf{C} \in M_{m \times 1}(F)$  is a single column  $\mathbf{x}$ , and  $\mathbf{R} \in M_{1 \times n}(F)$  is a single row, which we write as  $\mathbf{y}^\top$  with  $\mathbf{y} \in F^n$ . So  $\mathbf{A} = \mathbf{xy}^\top$ . Both are nonzero: the columns of  $\mathbf{C}$  and the rows of  $\mathbf{R}$  are linearly independent by [Theorem 4.3.1](#), and a list containing a zero vector is dependent. (This recovers one direction of Exercise (C2: Matrices of rank one).)
- Write  $\mathbf{A} = \mathbf{CR}$  with columns  $\mathbf{c}_1, \dots, \mathbf{c}_r$  of  $\mathbf{C}$  and rows  $\mathbf{r}_1, \dots, \mathbf{r}_r$  of  $\mathbf{R}$ , and let  $c_{ik}, r_{kl}$  be the entries of  $\mathbf{C}, \mathbf{R}$ . The  $m \times n$  matrix  $\mathbf{c}_k \mathbf{r}_k$  has  $(i, l)$ -entry  $c_{ik} r_{kl}$ , so by Theorem 0.9.2

(1),

$$(\mathbf{CR})_{il} = \sum_{k=1}^r c_{ik} r_{kl} = \sum_{k=1}^r (\mathbf{c}_k \mathbf{r}_k)_{il}, \quad \text{that is,} \quad \mathbf{A} = \mathbf{c}_1 \mathbf{r}_1 + \cdots + \mathbf{c}_r \mathbf{r}_r.$$

Each  $\mathbf{c}_k \mathbf{r}_k$  has rank 1 by the ( $\Leftarrow$ ) direction of Exercise (C2: Matrices of rank one), since  $\mathbf{c}_k \neq \mathbf{0}$  (a column of  $\mathbf{C}$ , whose columns are independent) and  $\mathbf{r}_k \neq \mathbf{0}$  (a row of  $\mathbf{R}$ , whose rows are independent). If  $\mathbf{A} = \mathbf{A}_1 + \cdots + \mathbf{A}_s$  with each  $\text{rank } \mathbf{A}_i = 1$ , then by Exercise (C1: Rank of a sum) applied repeatedly,  $r = \text{rank } \mathbf{A} \leq \text{rank } \mathbf{A}_1 + \cdots + \text{rank } \mathbf{A}_s = s$ .

(c) By [Example \(A column–row factorization\)](#),  $\mathbf{A} = \mathbf{CR}$  with  $\mathbf{C} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}$  and  $\mathbf{R} =$

$$\begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{pmatrix}, \text{ so}$$

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 & 3 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 0 & 3 \\ 2 & 4 & 0 & 6 \\ 3 & 6 & 0 & 9 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 1 & -2 \end{pmatrix}. \end{aligned}$$

#### 4.3.4.3 C. Going deeper

##### Exercise. C1: Frobenius's rank inequality

Let  $\mathbf{A} \in M_{m \times n}(F)$ ,  $\mathbf{B} \in M_{n \times p}(F)$ , and  $\mathbf{C} \in M_{p \times q}(F)$ . Write  $N = \text{nul}(\mathbf{A}) \subseteq F^n$ , and let  $S: F^n \rightarrow F^m$  be  $S\mathbf{y} = \mathbf{A}\mathbf{y}$ .

- Show that  $\text{col}(\mathbf{AB}) = \{\mathbf{A}\mathbf{y} : \mathbf{y} \in \text{col}(\mathbf{B})\}$ , and deduce from Rank–Nullity for  $S|_{\text{col}(\mathbf{B})}$  that  $\text{rank}(\mathbf{AB}) = \text{rank } \mathbf{B} - \dim(N \cap \text{col}(\mathbf{B}))$ .
- Show similarly that  $\text{rank}(\mathbf{ABC}) = \text{rank}(\mathbf{BC}) - \dim(N \cap \text{col}(\mathbf{BC}))$ .
- Prove that  $\text{col}(\mathbf{BC}) \subseteq \text{col}(\mathbf{B})$ , and deduce **Frobenius's inequality**

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

- Deduce Sylvester's inequality for matrices from (c).

*Hint: for (d), look for a special choice of the middle factor.*

*Solution.*

- By Theorem 0.9.2, every vector  $\mathbf{ABx}$  equals  $\mathbf{Ay}$  with  $\mathbf{y} = \mathbf{Bx} \in \text{col}(\mathbf{B})$ ; conversely every  $\mathbf{y} \in \text{col}(\mathbf{B})$  is  $\mathbf{Bx}$  for some  $\mathbf{x}$ , and then  $\mathbf{Ay} = (\mathbf{AB})\mathbf{x} \in \text{col}(\mathbf{AB})$ . So  $\text{col}(\mathbf{AB}) = \text{im}(S|_{\text{col}(\mathbf{B})})$ . The kernel of  $S|_{\text{col}(\mathbf{B})}$  is  $N \cap \text{col}(\mathbf{B})$ , and  $\text{col}(\mathbf{B})$  is finite-dimensional, so by Theorem 2.3.1,

$$\text{rank } \mathbf{B} = \dim \text{col}(\mathbf{B}) = \text{rank}(\mathbf{AB}) + \dim(N \cap \text{col}(\mathbf{B})).$$

- (b) The same argument with  $\mathbf{BC}$  in place of  $\mathbf{B}$  gives  $\text{col}(\mathbf{ABC}) = \text{im}(S|_{\text{col}(\mathbf{BC})})$ , with kernel  $N \cap \text{col}(\mathbf{BC})$ , so  $\text{rank}(\mathbf{BC}) = \text{rank}(\mathbf{ABC}) + \dim(N \cap \text{col}(\mathbf{BC}))$ .
- (c) Every vector  $\mathbf{BCx} = \mathbf{B}(\mathbf{Cx})$  lies in  $\text{col}(\mathbf{B})$ , so  $\text{col}(\mathbf{BC}) \subseteq \text{col}(\mathbf{B})$ . Hence  $N \cap \text{col}(\mathbf{BC}) \subseteq N \cap \text{col}(\mathbf{B})$ , and by Theorem 1.6.7,  $\dim(N \cap \text{col}(\mathbf{BC})) \leq \dim(N \cap \text{col}(\mathbf{B}))$ . By (b) and then (a),

$$\begin{aligned} \text{rank}(\mathbf{ABC}) &= \text{rank}(\mathbf{BC}) - \dim(N \cap \text{col}(\mathbf{BC})) \\ &\geq \text{rank}(\mathbf{BC}) - \dim(N \cap \text{col}(\mathbf{B})) \\ &= \text{rank}(\mathbf{BC}) - \text{rank } \mathbf{B} + \text{rank}(\mathbf{AB}). \end{aligned}$$

Rearranging gives Frobenius's inequality.

- (d) Let  $\mathbf{A} \in M_{m \times n}(F)$  and  $\mathbf{C} \in M_{n \times q}(F)$ , and apply (c) with  $p = n$  and  $\mathbf{B} = \mathbf{I}_n$ . Then  $\text{rank}(\mathbf{AC}) + n \geq \text{rank } \mathbf{A} + \text{rank } \mathbf{C}$ , since  $\text{rank } \mathbf{I}_n = n$ . This is Theorem 4.3.5 for matrices.

### Exercise. C2: Counting over $\mathbb{F}_2$

- (a) Show that  $M_2(\mathbb{F}_2)$  contains exactly 1 matrix of rank 0, 6 of rank 2, and 9 of rank 1.
- (b) Deduce the sizes of the equivalence classes of  $M_2(\mathbb{F}_2)$ .
- (c) Compare with similarity: how many matrices of  $M_2(\mathbb{F}_2)$  are similar to  $\mathbf{I}_2$ ? Is the class of  $\mathbf{I}_2$  under equivalence a union of several similarity classes?

*Solution.*

- (a) The only matrix of rank 0 is  $\mathbf{O}$ . A matrix has rank 2 exactly when its two columns span  $\mathbb{F}_2^2$ , that is (Theorem 1.6.4), when they form a basis of  $\mathbb{F}_2^2$ . The first column can be any of the 3 nonzero vectors of  $\mathbb{F}_2^2$ . The second must lie outside the span of the first, which has 2 elements, so there are  $4 - 2 = 2$  choices. That gives  $3 \cdot 2 = 6$  matrices of rank 2. Since  $|M_2(\mathbb{F}_2)| = 2^4 = 16$ , the remaining  $16 - 1 - 6 = 9$  matrices have rank 1.
- (b) By Theorem 4.3.3 (c), the classes are the sets of matrices of a fixed rank, of sizes 1, 9, and 6.
- (c)  $\mathbf{I}_2$  is similar only to itself, so its similarity class has 1 element, while its equivalence class, the invertible matrices, has 6. Since similar matrices are equivalent, each equivalence class is a union of similarity classes, and here the class of the 6 invertible matrices contains the similarity class  $\{\mathbf{I}_2\}$  together with others. For example  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  is invertible but not similar to  $\mathbf{I}_2$ , so there are at least two similarity classes inside.