

# Linear Algebra

A summary for MIT 18.06SC

Jiaxin Hu

June 30, 2020

## 1 Matrices & Spaces

### 1.1 Basic concepts

- Given vectors  $v_1, \dots, v_n$  and scalars  $c_1, \dots, c_n$ , the sum  $c_1v_1 + \dots + c_nv_n$  is called a *linear combination* of  $v_1, \dots, v_n$ .  
 $c_1=c_2=\dots=c_n=0$
- The vectors  $v_1, \dots, v_n$  are *linearly independent* (or just *independent*) if  $c_1v_1 + \dots + c_nv_n = 0$  holds only when all  $c_1, \dots, c_n = 0$ . If the vectors  $v_1, \dots, v_n$  are *dependent*, there exist scalars  $c_1, \dots, c_n$  which are not all equal to 0 such that  $c_1v_1 + \dots + c_nv_n = 0$ .
- Given a matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  and a vector  $x \in \mathbb{R}^n$ , the multiplication  $\mathbf{A}x$  is a linear combination of the columns of  $\mathbf{A}$ , and  $x^T \mathbf{A}$  is a linear combination of the rows of  $\mathbf{A}$ .
- Matrix multiplication is typically not communicative, i.e.  $\mathbf{AB} \neq \mathbf{BA}$ . Lemma 1 describes a special case where matrix multiplication is communicative.
- Suppose  $\mathbf{A}$  is a square matrix. The matrix  $\mathbf{A}$  is *invertible* or *non-singular* if there exists an  $\mathbf{A}^{-1}$  such that  $\mathbf{A}^{-1}\mathbf{A} = \mathbf{AA}^{-1} = \mathbf{I}$ . Otherwise, the matrix  $\mathbf{A}$  is *singular*, and the determinant of  $\mathbf{A}$  is 0.
- The inverse of a matrix product  $\mathbf{AB}$  is  $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ . The product of invertible matrices is still invertible.
- The transpose of a matrix product  $\mathbf{AB}$  is  $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ . For any invertible matrix  $\mathbf{A}$ ,  $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$ .
- A matrix  $\mathbf{Q}$  is orthogonal if  $\mathbf{Q}^T = \mathbf{Q}^{-1}$ . A matrix  $\mathbf{Q}$  is unitary if  $\mathbf{Q}^* = \mathbf{Q}^{-1}$ , where  $\mathbf{Q}^*$  is the conjugate transpose of  $\mathbf{Q}$ .

**Lemma 1** (Communicative matrix multiplication). *For matrices  $\mathbf{A} \in \mathbb{R}^{n \times n}$  and  $\mathbf{B} \in \mathbb{R}^{n \times n}$ , the matrix multiplication of  $\mathbf{A}$  and  $\mathbf{B}$  is communicative, i.e.  $\mathbf{AB} = \mathbf{BA}$ , if  $\mathbf{A}$  and  $\mathbf{B}$  have the same set of eigenvectors corresponding to their non-zero eigenvalues.*

*Proof.* If  $\mathbf{A}$  and  $\mathbf{B}$  have the same set of eigenvectors corresponding to their non-zero eigenvalues, there exists a matrix  $\mathbf{Q} \in \mathbb{R}^{n \times n}$  such that

$$\mathbf{A} = \mathbf{Q}\mathbf{D}_A\mathbf{Q}^{-1}, \quad \mathbf{B} = \mathbf{Q}\mathbf{D}_B\mathbf{Q}^{-1},$$

where columns of  $\mathbf{Q}$  are eigenvectors of  $\mathbf{A}$  and  $\mathbf{B}$ , and  $\mathbf{D}_A \in \mathbb{R}^{n \times n}, \mathbf{D}_B \in \mathbb{R}^{n \times n}$  are diagonal matrices whose diagonal elements are eigenvalues of  $\mathbf{A}$  and  $\mathbf{B}$ , respectively. Because matrix multiplication is communicative for two diagonal matrices with same dimensions, we have

$$\mathbf{AB} = \mathbf{Q}\mathbf{D}_A\mathbf{Q}^{-1}\mathbf{Q}\mathbf{D}_B\mathbf{Q}^{-1} = \mathbf{Q}\mathbf{D}_A\mathbf{D}_B\mathbf{Q}^{-1} = \mathbf{Q}\mathbf{D}_B\mathbf{D}_A\mathbf{Q}^{-1} = \mathbf{Q}\mathbf{D}_B\mathbf{Q}^{-1}\mathbf{Q}\mathbf{D}_A\mathbf{Q}^{-1} = \mathbf{BA}.$$

Therefore, the matrix multiplication of  $\mathbf{A}$  and  $\mathbf{B}$  is communicative.  $\square$

## 1.2 Permutation of matrices

~~Let  $\mathbf{A}$  be a matrix. the rows of  $\mathbf{A}$~~  extra whitespace?  
~~For any matrix  $\mathbf{A}$ , we swap its rows by multiplying a permutation matrix  $\mathbf{P}$  on the left of  $\mathbf{A}$ . For example,~~

$$\mathbf{PA} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} a_3 \\ a_1 \\ a_2 \end{bmatrix}$$

where  $a_k$  refers to the  $k$ -th row of  $\mathbf{A}$ . The inverse of the permutation matrix  $\mathbf{P}$  is  $\mathbf{P}^{-1} = \mathbf{P}^T$ , which implies the orthogonality of permutation matrix. For an  $n \times m$  matrix, there are  $n!$  different row permutation matrices, which form a *multiplicative group*.

Similarly, we also swap the columns of the matrix  $\mathbf{A}$  by multiplying a permutation matrix on the right of  $\mathbf{A}$ .

## 1.3 Elimination of matrices

Elimination is an important technique in linear algebra. We eliminate the matrix by multiplications and subtractions. Take a 3-by-3 matrix  $\mathbf{A}$  as an example.

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} \xrightarrow{\text{step 1}} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix} \xrightarrow{\text{step 2}} \mathbf{U} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 5 \end{bmatrix}$$

In step 1, we choose the number 1 in row 1 column 1 as a *pivot*, then we recopy the first row, multiply an appropriate number (in this case, 3) and subtract those values from the numbers in the second row. We have thus eliminated 3 in row 2 column 1. Similarly, in step 2, we choose 2 in row 2 column 2 as a pivot and eliminate the number 4 in row 3 column 2. The number 5 in row 3 column 3 is also a pivot. The matrix  $\mathbf{U}$  is an upper triangular matrix.

The *elimination matrix* used to eliminate the entry in row  $m$  column  $n$  is denoted as  $\mathbf{E}_{mn}$ . In the previous example,

$$\mathbf{E}_{21}\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix}; \quad \mathbf{E}_{32}(\mathbf{E}_{21}\mathbf{A}) = \mathbf{U}.$$

Pivots are non-zero. If there ~~is~~ a 0 in the pivot position, we ~~must~~ exchange the row with one below to get a non-zero value in pivot position. If there ~~is not~~ non-zero value below the 0 pivot, we skip this column and find a pivot in the next column. ~~exist no non-zero values~~

We write  $\mathbf{E}_{32}(\mathbf{E}_{21}\mathbf{A}) = (\mathbf{E}_{32}\mathbf{E}_{21})\mathbf{A} = \mathbf{U}$ , because matrix multiplication is associative. Let  $\mathbf{E}$  denote the product of all elimination matrices. If we need to permute the rows during the process, we multiply a permutation matrix on the left of  $\mathbf{A}$ . Therefore, the elimination process of  $\mathbf{A}$  is

$$\mathbf{E}\mathbf{PA} = \mathbf{U}, \tag{1}$$

where  $\mathbf{U}$  is an upper triangular matrix.

Next, we prove the invertibility of the elimination matrix.

*Let  $\mathbf{E}$  in ... be an elimination matrix that ...*

**Lemma 2** (Invertibility of elimination matrix). *Suppose there is an elimination matrix  $\mathbf{E}_{ij} \in \mathbb{R}^{n \times n}$  that multiplies a scalar  $-c$  to the  $j$ -th row and subtracts the row from  $i$ -th row, where  $i \neq j$ . The matrix  $\mathbf{E}_{ij}$  is invertible.*

*Proof.* We write the elimination matrix as

$$\mathbf{E}_{ij} = \mathbf{I}_n + ce_i e_j^T,$$

where  $e_i \in \mathbb{R}^n$  denotes the vector with value 1 in the  $i$ -th entry and value 0 elsewhere. Because  $i \neq j$ ,  $e_i^T e_j = 0$ . We have

$$(\mathbf{I}_n + ce_i e_j^T)(\mathbf{I}_n - ce_i e_j^T) = \mathbf{I}_n - c^2 e_i e_j^T e_i e_j^T = \mathbf{I}_n; \quad (\mathbf{I}_n - ce_i e_j^T)(\mathbf{I}_n + ce_i e_j^T) = \mathbf{I}_n.$$

Therefore,  $\mathbf{I}_n - ce_i e_j^T$  is the inverse of  $\mathbf{E}_{ij}$ . The elimination matrix  $\mathbf{E}_{ij}$  is invertible.  $\square$

**Corollary 1** (Inverse of elimination matrix). Suppose the elimination matrix  $\mathbf{E}_{ij}$  in lemma 2 is a lower/upper-triangular matrix. The inverse  $\mathbf{E}_{ij}^{-1}$  is also a lower/upper-triangular matrix.

*Proof.* By the proof of lemma 2, the matrix  $\mathbf{E}_{ij}$  and its inverse are written as

explain acronym in its first appearance  $\mathbf{E}_{ij} = \mathbf{I}_n + ce_i e_j^T, \quad \mathbf{E}_{ij}^{-1} = \mathbf{I}_n - ce_i e_j^T.$

Without loss of generality (WLOG)

*WLOG*, we assume that  $\mathbf{E}_{ij}$  is a lower-triangular matrix. Then  $ce_i e_j^T$  and  $-ce_i e_j^T$  are also lower-triangular matrices. Therefore,  $\mathbf{E}_{ij}^{-1}$  is a lower-triangular matrix.  $\square$

## 1.4 Gauss-Jordan Elimination

We also use elimination to find the inverse of an invertible matrix.

*of  $\mathbf{A}$ , denoted  $\mathbf{A}^{-1}$ ,*

Suppose  $\mathbf{A} \in \mathbb{R}^{n \times n}$  is an invertible matrix. The inverse of  $\mathbf{A}$ ,  $\mathbf{A}^{-1}$ , satisfies

$$\mathbf{AA}^{-1} = \mathbf{I}_n. \quad \text{exists an elimination matrix sides} \quad (2)$$

Suppose there is an elimination  $\mathbf{E}$  such that  $\mathbf{EA} = \mathbf{I}_n$ . Multiplying  $\mathbf{E}$  on the both side of the equation (2), we have  $\mathbf{EAA}^{-1} = \mathbf{A}^{-1} = \mathbf{E}$ . To obtain a such  $\mathbf{E}$ , we eliminate the augmented matrix  $[\mathbf{A}|\mathbf{I}_n]$  until  $\mathbf{A}$  becomes  $\mathbf{I}_n$ . Then, the augmented matrix becomes  $\mathbf{E}[\mathbf{A}|\mathbf{I}_n] = [\mathbf{I}_n|\mathbf{E}]$ , where  $\mathbf{E}$  is the inverse of  $\mathbf{A}$ .

We call this elimination process of finding  $\mathbf{E}$  as *Gauss-Jordan Elimination*.

## 1.5 Factorization of matrices

*any square matrix  $\mathbf{A}$  has the factorization (1)*

By elimination, for any square matrix  $\mathbf{A}$ , we have equation (1). By lemma 2,  $\mathbf{E}$  is invertible. We multiply  $\mathbf{E}^{-1}$  on both sides of equation (1). We have, yields

*Multiplying*

$$\mathbf{PA} = \mathbf{E}^{-1}\mathbf{U}.$$

Note that  $\mathbf{E}$  is a lower-triangular matrix. By corollary 1,  $\mathbf{E}^{-1}$  is also a lower-triangular matrix. Let  $\mathbf{L}$  denote  $\mathbf{E}^{-1}$ , where the letter  $L$  refers to “lower triangular”. Therefore, any square matrix  $\mathbf{A}$  has a factorization:

$$\mathbf{PA} = \mathbf{LU}, \quad (3)$$

where  $\mathbf{U}$  is an upper triangular matrix with pivots on the diagonal,  $\mathbf{L}$  is a lower triangular matrix with ones on the diagonal, and  $\mathbf{P}$  is a permutation matrix. However, the equation (3) is not the unique factorization of  $\mathbf{A}$ . For example,  $c\mathbf{L}$  and  $c^{-1}\mathbf{U}$  also factorize  $\mathbf{A}$ , where  $c$  is a non-zero scalar.

## 1.6 Time complexity of elimination

~~a single elimination → a single elimination step~~

For an  $n$ -by- $n$  matrix, a single elimination multiplies a selected row and subtracts the selected row from another row. ~~A single elimination~~ requires  $\mathcal{O}(n)$  operations. To eliminate the elements below the first diagonal element, we need to repeat single elimination  $(n-1)$  times and thus require  $\mathcal{O}(n^2)$  operations. Similarly, we require  $\mathcal{O}((n-1)^2)$  operations to eliminate the elements below the second diagonal element. Repeat the elimination until we meet the  $n$ -th diagonal element. Therefore, we require total  $\mathcal{O}(n^3)$  operations to obtain an upper-triangular matrix by elimination:

in total

$$1^2 + 2^2 + \cdots + (n)^2 = \sum_i^n i^2 \approx \int_0^n x^2 dx = \frac{1}{3}n^3 = \mathcal{O}(n^3).$$

## 1.7 Reduced row echelon form of matrices

In previous sections, we convert a matrix  $\mathbf{A}$  to an upper triangular matrix  $\mathbf{U}$ . Next, we convert  $\mathbf{U}$  into the *reduced row echelon form* (RREF), which is a simpler form than upper triangle. We use  $\mathbf{R} = RREF(\mathbf{A})$  to denote the reduced row echelon form of  $\mathbf{A}$ . In  $\mathbf{R}$ , the pivots are equal to 1, and the elements above and below the pivots are eliminated to 0. In the previous example,

$$\mathbf{U} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 5 \end{bmatrix} \xrightarrow{\text{make pivots = 1}} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{\text{0 above and below pivots}} \mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

There is another example,

$$\mathbf{U} = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{make pivots = 1}} \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{0 above and below pivots}} \mathbf{R} = \begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Assume that there are  $r$  pivots in  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . With proper permutation, the matrix  $\mathbf{R}$  is in form  $\begin{bmatrix} \mathbf{I}_r & \mathbf{F} \\ 0 & 0 \end{bmatrix}$ , where  $\mathbf{F} \in \mathbb{R}^{r \times (n-r)}$  is an arbitrary matrix. The columns in  $\mathbf{A}$  which correspond to the identity matrix  $\mathbf{I}_r$  are called *pivot columns*. The other columns are *free columns*.

“,” and

## 1.8 Vector space, Subspace and Column space

~~closed~~

- *Vector space* is a collection of vectors that is ~~close~~ under linear combination (addition and multiplication by any real number); i.e. for any vectors in the collection, all the combinations of these vectors are still in the collection.
- *Subspaces of the vector space* is a vector space that is contained inside of another vector space.

Note that any vector space or subspace must include an origin. For a vector space  $\mathcal{A}$ , the subspace of  $\mathcal{A}$  can be  $\mathcal{A}$  itself or a set that contains only a zero vector.

~~consists of linear combinations~~

- Vectors  $v_1, \dots, v_n$  *span* a space that ~~consists~~ all the ~~combination~~ of these vectors.
- *Column space* of a matrix  $\mathbf{A}$  is the space spanned by the columns of  $\mathbf{A}$ . Let  $C(\mathbf{A})$  denote the column space of  $\mathbf{A}$ .

~~contains~~

If  $v_1, \dots, v_n$  span a space  $\mathcal{S}$ , then  $\mathcal{S}$  is the smallest space that ~~contain~~ these vectors.

- *Basis* of a vector space is ~~a sequence~~<sup>a set</sup> of vectors  $v_1, \dots, v_n$  that satisfy: (1)  $v_1, \dots, v_n$  are independent; (2)  $v_1, \dots, v_n$  span the space.
- *Dimension* of a space is the number of vectors in a basis of this space. Let  $\dim(\mathcal{A})$  denote the dimension of space  $\mathcal{A}$ .

## 1.9 Matrix rank

Let  $\text{rank}(\mathbf{A})$  denote the *rank* of matrix  $\mathbf{A}$ . We have

$$\text{rank}(\mathbf{A}) \stackrel{\Delta}{=} \dim(C(\mathbf{A})) = \# \text{ of pivot columns of } \mathbf{A}. \quad (4)$$

If  $\mathbf{A} \in \mathbb{R}^{m \times n}$  and  $\text{rank}(\mathbf{A}) = r$ , we have  $r \leq \min\{m, n\}$ . The matrix is *full rank* if  $r = \min\{m, n\}$ .

The rank of a square matrix is closely related to the invertibility.

**Lemma 3** (Full rankness and invertibility). *A square matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  is full rank, if and only if  $\mathbf{A}$  is an invertible matrix.*

*Proof.* ( $\Rightarrow$ ) Suppose  $\mathbf{A}$  is full rank. (Good. The current structure is better than your previous one)

Let  $\mathbf{R}$  denote the RREF form of  $\mathbf{A}$ . There exist an elimination matrix  $\mathbf{E}$  and a permutation matrix  $\mathbf{P}$  such that

$$\mathbf{E}\mathbf{P}\mathbf{A} = \mathbf{R}.$$

By the full rankness of  $\mathbf{A}$ ,  $\mathbf{A}$  has  $n$  pivot columns, and thus  $\mathbf{R} = \mathbf{I}_n$ . By lemma 2,  $\mathbf{E}$  is invertible. The permutation matrix  $\mathbf{P}$  is also invertible. Then, the matrix product  $\mathbf{EP}$  is invertible, and  $\mathbf{A}$  is the inverse of  $\mathbf{EP}$ . Therefore,  $\mathbf{A}$  is invertible.

( $\Leftarrow$ ) Suppose  $\mathbf{A}$  is invertible. We prove the full rankness of  $\mathbf{A}$  by contradiction.

Assume that  $\text{rank}(\mathbf{A}) < n$ . By equation (4),  $\dim(C(\mathbf{A})) = \text{rank}(\mathbf{A}) < n$ , which implies that the columns of  $\mathbf{A}$  are linearly dependent. Then, there exists a non-zero vector  $v$  such that

$$\mathbf{A}v = 0. \quad (5)$$

By assumption, the inverse  $\mathbf{A}^{-1}$  exists. Multiplying  $\mathbf{A}^{-1}$  on both sides of equation (5), we have

$$v = \mathbf{A}^{-1}0 = 0. \quad (6)$$

However, equation (6) contradicts the fact that  $v$  is a non-zero vector. Therefore,  $\mathbf{A}$  is full rank.  $\square$

The rank of  $\mathbf{A}$  also affects the number of solutions to the system  $\mathbf{Ax} = b$ . We will discuss the relationship between matrix rank and the solutions in next section.

## 2 Solving $\mathbf{Ax} = b$

In this section, we discuss the solutions to the linear system  $\mathbf{Ax} = b$ , where  $\mathbf{A} \in \mathbb{R}^{m \times n}$  is a matrix, and  $x \in \mathbb{R}^n$ ,  $b \in \mathbb{R}^m$  are vectors.

## 2.1 Solving $Ax = 0$ : Nullspace

The *nullspace* of matrix  $\mathbf{A}$  is the collection of all solutions  $x$  to the system  $\mathbf{Ax} = 0$ . Let  $N(\mathbf{A})$  denote the nullspace of  $\mathbf{A}$ .

**Lemma 4** (Nullspace). *The nullspace of matrix  $\mathbf{A}$  is a vector space.*

*Proof.* We only need to prove that  $N(\mathbf{A})$  is closed under linear combination. For  $\forall v_1, v_2 \in N(\mathbf{A})$ , we have,

$$\mathbf{A}(c_1v_1 + c_2v_2) = c_1\mathbf{Av}_1 + c_2\mathbf{Av}_2 = 0, \quad \forall c_1, c_2 \in \mathbb{R}. \quad (7)$$

The equation (7) implies that  $N(\mathbf{A})$  is closed under linear combination. Therefore,  $N(\mathbf{A})$  is a vector space.  $\square$

**Lemma 5** (The rank of nullspace). *If  $\text{rank}(\mathbf{A}) = r$ , the rank of nullspace  $\text{rank}(N(\mathbf{A})) = n - r$ .*

*Proof.* Let  $\mathbf{R}$  denote the RREF( $\mathbf{A}$ ). We write  $\mathbf{R}$  in form  $\mathbf{R} = \begin{bmatrix} \mathbf{I}_r & \mathbf{F} \\ 0 & 0 \end{bmatrix}$ , where  $\mathbf{F} \in \mathbb{R}^{r \times (n-r)}$  is

an arbitrary matrix. Let  $\mathbf{X} = \begin{bmatrix} -\mathbf{F} \\ \mathbf{I}_{n-r} \end{bmatrix}$ . We have

Is ``a special solution'' a terminology?

If so, then say "is called a special solution".

If not, then fine.

$$\mathbf{RX} = \begin{bmatrix} \mathbf{I}_r & \mathbf{F} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -\mathbf{F} \\ \mathbf{I}_{n-r} \end{bmatrix} = 0.$$

Therefore, each column of  $\mathbf{X}$  is a special solution to the system  $\mathbf{Ax} = 0$ . Next, we show that any vector  $x \in N(\mathbf{A})$  is a linear combination of the columns of  $\mathbf{X}$ .

Suppose there is a solution  $x = (x_1, x_2) \in N(\mathbf{A})$ , where  $x_1 \in \mathbb{R}^r$  and  $x_2 \in \mathbb{R}^{n-r}$ . We have

$$\mathbf{Rx} = \begin{bmatrix} \mathbf{I}_r & \mathbf{F} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + \mathbf{Fx}_2 \\ 0 \end{bmatrix} = 0.$$

This implies that  $x_1 = -\mathbf{Fx}_2$ , and  $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\mathbf{F} \\ \mathbf{I}_{n-r} \end{bmatrix} x_2 = \mathbf{X}x_2$ . Any vector in  $N(\mathbf{A})$  is a linear combination of the columns of  $\mathbf{X}$ , i.e.  $C(\mathbf{X}) = N(\mathbf{A})$ . Therefore, the rank of nullspace  $\text{rank}(N(\mathbf{A})) = \dim(C(\mathbf{X})) = n - r$ . type "i.e.\ C(X)=..." in latex. Note "\\" after ":"  $\square$

Recall the definitions of pivot columns and free columns. In  $\mathbf{Ax} = b$ , the variables in  $x$  that correspond to pivot columns are called *pivot variables*, and others are *free variables*. If  $\text{rank}(\mathbf{A}) = r$ , there are  $n - r$  free variables.

In the proof of lemma 5, the columns of  $\mathbf{X} = \begin{bmatrix} -\mathbf{F} \\ \mathbf{I}_{n-r} \end{bmatrix}$  are special solutions that compose the basis of  $N(\mathbf{A})$ . Practically, we find these special solutions by assigning 1 to a free variable and 0 to other free variables, and then we solve the system  $\mathbf{Ax} = 0$ .

## 2.2 Solving $Ax = b$ : complete solutions

**Lemma 6** (Solvability of  $\mathbf{Ax} = b$ ). *The system  $\mathbf{Ax} = b$  is solvable only when  $b \in C(\mathbf{A})$ .*

*Proof.* If  $\mathbf{Ax} = b$  is solvable, there exists a  $x$  such that  $\mathbf{Ax} = b$ . For any  $x$ ,  $\mathbf{Ax} \in C(\mathbf{A})$ . Therefore,  $b \in C(\mathbf{A})$ .  $\square$

**Lemma 7** (Complete solution). *The complete solution of  $\mathbf{Ax} = b$  is given by  $x_{\text{comp}} = x_p + x_n$ , where  $x_p$  is a particular solution such that  $\mathbf{Ax}_p = b$ , and  $x_n \in N(\mathbf{A})$ .*

*Proof.* Suppose  $x = x_p + x_0$  is ~~an arbitrary~~<sup>a</sup> solution to  $\mathbf{A}x = b$ . We have

$$\mathbf{A}x - \mathbf{A}x_p = \mathbf{A}(x - x_p) = \mathbf{A}x_0 = 0.$$

Therefore,  $x_0 \in N(\mathbf{A})$ . □

**what is the difference between “a particular solution” vs. “a special solution”? If they are the same, then stick with one phrase**

Usually, we find a particular solution by assigning 0 to free variables, and we solve the system

$$\mathbf{A}x = b.$$

The following table discusses the rank of  $\mathbf{A}$ , the form of  $\mathbf{R}$ , the dimension of nullspace  $N(\mathbf{A})$ , and the number of solutions to  $\mathbf{A}x = b$ .

	$r = m = n$	$r = n < m$	$r = m < n$	$r < m, r < n$
$\mathbf{R}$	$\mathbf{I}$	$\begin{bmatrix} \mathbf{I} \\ 0 \end{bmatrix}$	$[\mathbf{I} \quad \mathbf{F}]$	$\begin{bmatrix} \mathbf{I} & \mathbf{F} \\ 0 & 0 \end{bmatrix}$
$\dim(N(\mathbf{A}))$	0	0	$n - r$	$n - r$
# solutions to $\mathbf{A}X = b$	1	0 or 1	infinitely many	0 or infinitely many