Linear Algebra

A summary for MIT 18.06SC

Jiaxin Hu

June 27, 2020

1 Matrices & Spaces

1.1 Basic concepts

"a" linear combination

- Given vectors $v_1, ..., v_n$ and scalars $c_1, ..., c_n$, the sum $c_1 v_1 + \cdots + c_n v_n$ is called the *linear combination* of $v_1, ..., v_n$.

space before bracket

- The vectors $v_1, ..., v_n$ are *linearly independent*(or just *indepedent*) if $c_1 v_1 + \cdots + c_n v_n = 0$ holds only when all $c_1, ..., c_n = 0$. If the vectors $v_1, ..., v_n$ are *dependent*, there exist scalars $c_1, ..., c_n$ which are not all equal to 0 ~~and satisfy~~ $c_1 v_1 + \cdots + c_n v_n = 0$.

such that

- Given a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and a vector $x \in \mathbb{R}^n$, the multiplication $\boldsymbol{A}x$ is a linear combination of the columns of $\boldsymbol{A}$ and $x^T \boldsymbol{A}$ is a linear combination of the rows of $\boldsymbol{A}$.

- Matrix multiplication is not communicative, i.e. $\boldsymbol{AB} \neq \boldsymbol{BA}$. Q: When is matrix multiplication commutative? Hint: characterize the conditions in terms of spectral property.

- Suppose $\boldsymbol{A}$ is a square matrix. The matrix $\boldsymbol{A}$ is *invertible* or *non-singular* if there exists a $\boldsymbol{A}^{-1}$ such that $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-1} = I$. Otherwise, the matrix $\boldsymbol{A}$ is singular, i.e. its determinant is 0 and does not have inverse matrix.

ambiguous subject

- The inverse of a matrix product $\boldsymbol{AB}$ is $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$. The product of invertible matrices is still invertible.

- The transpose of a matrix product $\boldsymbol{AB}$ is $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$. For any invertible matrix $\boldsymbol{A}$, $(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T$.

- A matrix $\boldsymbol{Q}$ is orthogonal if $\boldsymbol{Q}^T = \boldsymbol{Q}^{-1}$. A matrix $\boldsymbol{Q}$ is unitary if $\boldsymbol{Q}^* = \boldsymbol{Q}^{-1}$, where $\boldsymbol{Q}^*$ is the conjugate transpose of $\boldsymbol{Q}$.

1.2 Permutation of matrices

For any matrix $\boldsymbol{A}$, we can swap its rows by multiplying a *permutation matrix* $\boldsymbol{P}$ on the left of $\boldsymbol{A}$. For example,

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

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

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

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 $\boldsymbol{A}$ as example.
an example

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

In step 1, we choose the number 1 in row 1 column 1 as a *pivot*, then we recopy the first row and 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 $\boldsymbol{U}$ is an upper traingular matrix.

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

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

Pivots can not be 0. 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, then we skip this column and find a pivot in next column.

Since matrix multiplication is associative, we can write $\boldsymbol{E}_{32}(\boldsymbol{E}_{21}\boldsymbol{A}) = (\boldsymbol{E}_{32}\boldsymbol{E}_{21})\boldsymbol{A} = \boldsymbol{U}$. Let $\boldsymbol{E}$ denote the product of all elimination matrices. If we need to permute the rows during the process, we have $\boldsymbol{EPA} = \boldsymbol{U}$, where $\boldsymbol{P}$ is the product of all the needed permutation matrices.

We also prove the invertibility of the elimination matrix.

multiplies subtracts
Lemma 1 (Invertiblity of elimination matrix)**.** *Suppose there is an elimination matrix $\boldsymbol{E}_{ij} \in \mathbb{R}^{n\times n}$ that ~~means multiplying~~ a scalar $-c$ to the j-th row and ~~subtracting~~ the row from i-th row, where $i \neq j$. Then, $\boldsymbol{E}_{ij}$ is invertible.*

Proof. The elimination matrix can be written as:

$$\boldsymbol{E}_{ij} = \boldsymbol{I}_n + ce_ie_j^T,$$

(missing article) no such term called “identity vector”. “basis vector”
where $e_i \in \mathbb{R}^n$ is identity vector with value 1 on the i-th entry and value 0 on the other entries. Note that $e_i^Te_j = 0$ because $i \neq j$. Therefore, we have

$$(\boldsymbol{I}_n + ce_ie_j^T)(\boldsymbol{I}_n - ce_ie_j^T) = \boldsymbol{I}_n - c^2e_ie_j^Te_ie_j^T = \boldsymbol{I}_n; \quad (\boldsymbol{I}_n - ce_ie_j^T)(\boldsymbol{I}_n + ce_ie_j^T) = \boldsymbol{I}_n$$ punctuation

Thus, $\boldsymbol{I}_n - ce_ie_j^T$ is the inverse of $\boldsymbol{E}_{ij}$ and $\boldsymbol{E}_{ij}$ is invertible. □

Highlighted sentences have either wrong or bad grammar.

1.4 Gauss-Jordan Elimination

Consider an invertible matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, one of the effective ways to find the inverse of $\boldsymbol{A}$ uses elimination.

The inverse of $\boldsymbol{A}$, $\boldsymbol{A}^{-1}$, satisfies $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}_n$. Suppose there is an elimination $\boldsymbol{E}$ such that $\boldsymbol{E}\boldsymbol{A} = \boldsymbol{I}_n$. Multiplying $\boldsymbol{E}$ on the both side of the equation, we have $\boldsymbol{E}\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1} = \boldsymbol{E}$. To obtain an such $\boldsymbol{E}$, we do elimination to the *augmented matrix* $[\boldsymbol{A}|\boldsymbol{I}_n]$ until $\boldsymbol{A}$ becomes $\boldsymbol{I}_n$.

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

1.5 Factorization of matrices

you should type ``.... '' in latex

By elimination, for any square matrix $\boldsymbol{A}$, we have $\boldsymbol{E}\boldsymbol{P}\boldsymbol{A} = \boldsymbol{U}$ where $\boldsymbol{U}$ is an upper triangular matrix. By "canceling" the elimination matrix $\boldsymbol{E}$, we get $\boldsymbol{P}\boldsymbol{A} = \boldsymbol{E}^{-1}\boldsymbol{U}$. Because $\boldsymbol{E}$ is invertible, the inverse $\boldsymbol{E}^{-1}$ exists. Note that $\boldsymbol{E}$ is lower triangular matrix, the inverse of $\boldsymbol{E}$ is also a lower triangular matrix. We use $\boldsymbol{L}$ to denote $\boldsymbol{E}^{-1}$. Therefore, we can decompose an arbitrary square matrix $\boldsymbol{A}$ as:

$$\boldsymbol{P}\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U},$$

where $\boldsymbol{U}$ is an upper triangular matrix with pivots on the diagonal, $\boldsymbol{L}$ is lower triangular matrix with ones on the diagonal, and $\boldsymbol{P}$ is a permutation matrix. Note that, there may exist other ways to decompose $\boldsymbol{P}\boldsymbol{A}$ into $\boldsymbol{L}\boldsymbol{U}$, where $\boldsymbol{U}$ and $\boldsymbol{L}$ have different settings.

1.6 Time complexity of elimination

ambiguity.. see page 7 in phd-writing-Cochrane.pdf

For an n-by-n matrix, multiplying one row and then subtracting it from another row require $2n$ operations. There are n rows in the matrix, so the total number of operations used in elimination the first column is $2n^2$. The second row and column are shorter, which may cost $2(n-1)^2$ operations and so on. Therefore, the time complexity to factorize $\boldsymbol{A}$ into $\boldsymbol{L}\boldsymbol{U}$ is about $\mathcal{O}(n^3)$:

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

1.7 Reduced row echelon form of matrices

By continuing to use the method of elimination, we can convert $\boldsymbol{U}$ to a matrix $\boldsymbol{R}$ in reduced row echelon form, with pivots equal to 1 and zeros above and below the pivots. The matrix $\boldsymbol{R}$ is called the *reduced row echelon form of matrices*(RREF) of $\boldsymbol{A}$. In previous example,

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

For an another example,

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

With proper permutation, the matrix $\boldsymbol{R}$ can be written in form $[\boldsymbol{I} \quad \boldsymbol{F}]$, $\begin{bmatrix} \boldsymbol{I} & \boldsymbol{F} \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} \boldsymbol{I} \\ 0 \end{bmatrix}$, or just $\boldsymbol{I}$,

is an ambiguous

where $\boldsymbol{F}$ ~~can be~~ arbitrary matrix in proper dimension. The columns in $\boldsymbol{A}$ which correspond to the identity matrix $\boldsymbol{I}$ are called *pivot columns* and the other columns are *free columns*.

1.8 Vector space, Subspace and Column space

- *Vector space* is a collection of vectors that are closed 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 can only contain a zero vector. “only” should immediately precede the noun that being modified

grammar mistake I have pointed out before (during NeurIPS submission)

- Vectors $v_1, ..., v_n$ *span* a space that consists all ~~the~~ combination of those vectors.

- *Column space* of matrix $\boldsymbol{A}$ is the space spanned by ~~the~~ columns of $\boldsymbol{A}$. Let $C(\boldsymbol{A})$ denote the column space of $\boldsymbol{A}$.

Note that if $v_1, ..., v_n$ span a space $\mathcal{S}$, then $\mathcal{S}$ is the smallest space that contain those vectors.

- *Basis* of a vector space is a sequence of vectors $v_1, ..., v_n$ satisfies: (1) $v_1, ..., v_n$ are independent; (2) $v_1, ..., v_n$ span the space.

- *Dimension* of the space is the number of vectors in a basis of the space.

1.9 Matrix rank

The *rank* of a matrix $\boldsymbol{A}$ is defined as the dimension of the columns space of $\boldsymbol{A}$. Rank is also equal to the number of pivot columns of $\boldsymbol{A}$. That means:

$$\text{rank}(\boldsymbol{A}) = \#\text{ of pivot columns of } \boldsymbol{A} = \text{dimension of } C(\boldsymbol{A}).$$

We use r to denote the rank of $\boldsymbol{A}$. If $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, then we have $r \leq m, r \leq n$. We say the matrix is *full rank* if $r = n$ or $r = m$.

could you put this two inequality in one?

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

Lemma 2 (Full rankness and invertibility)**.** *A square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is full rank, if and only if $\boldsymbol{A}$ is an invertible matrix.*

Avoid using ``can” in formal writing. Please correct throughout the note.

Proof. First, assume $\boldsymbol{A}$ is full rank, we prove that $\boldsymbol{A}$ has an inverse. We can get find a RREF($\boldsymbol{A}$) by eliminations and permutations. There are $\boldsymbol{E}$ and $\boldsymbol{P}$ such that

$$\boldsymbol{EPA} = \boldsymbol{R}.$$

Since $\boldsymbol{A}$ is full rank, $\boldsymbol{A}$ have n pivots columns and that implies $\boldsymbol{R} = \boldsymbol{I}_n$. By lemma 1, $\boldsymbol{E}$ is invertible. The permutation matrix $\boldsymbol{P}$ is also invertible. Therefore, $\boldsymbol{EP}$ is invertible and $\boldsymbol{AEP} = \boldsymbol{I}_n$. This implies $\boldsymbol{A}$ is invertible.

Second, assume $\boldsymbol{A}$ has an inverse, we prove that $\boldsymbol{A}$ is full rank. We show this by contradiction. Assume $\boldsymbol{A}$ has an inverse $\boldsymbol{A}^{-1}$ and $\boldsymbol{A}$ is not full rank. Because the rank of $\boldsymbol{A}$ is equal to the dimension of $C(\boldsymbol{A})$, the columns of $\boldsymbol{A}$ are linearly dependent ~~without full rankness~~. So, there exist a non-zero vector v such that

(1) delete unnecessary logit terms such as “So”. (2) correct grammar mistakes

$$\boldsymbol{A}v = 0.$$

Multiplying $\boldsymbol{A}^{-1}$ on the both sides of the equation, we have

$$v = \boldsymbol{A}^{-1}0 = 0.$$

However, it contradicts to non-zeroness of v. Therefore, $\boldsymbol{A}$ must be full rank. □

The rank of $\boldsymbol{A}$ also effects the number of solutions to the system $\boldsymbol{A}x = b$. We will discuss it in next section.

2 Solving $Ax = b$

Here we discuss the solution situation of the linear system $Ax = b$. Without specific explanation, the matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and the vector $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$.

2.1 Solving $Ax = 0$: Nullspace

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

Lemma 3 (Nullspace)**.** *The nullspace of matrix $\boldsymbol{A}$ is a vector space.*

Proof. To show the $N(\boldsymbol{A})$ is a vector space, we need to show $N(\boldsymbol{A})$ is close to linear combination. For any integer k, take arbitrary vectors $v_1, ..., v_k \in N(\boldsymbol{A})$ and arbitrary scalars $c_1, ..., c_k$. We have,

$$\boldsymbol{A}(c_1 v_1 + \cdots + c_k v_k) = c_1 \boldsymbol{A} v_1 + \cdots + c_k \boldsymbol{A} v_k = 0.$$

Therefore, the linear combination $(c_1 v_1 + \cdots + c_k v_k) \in N(\boldsymbol{A})$. Then $N(\boldsymbol{A})$ is a vector space. □

Lemma 4 (The rank of nullspace)**.** *Suppose the rank of $\boldsymbol{A}$ is r, then the rank of $N(\boldsymbol{A})$ is $n - r$.*

Proof. Let $\boldsymbol{R}$ denote the RREF($\boldsymbol{A}$). The matrix $\boldsymbol{R}$ can be written in form $\boldsymbol{R} = \begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{F} \\ 0 & 0 \end{bmatrix}$ where $\boldsymbol{F} \in \mathbb{R}^{r \times (n-r)}$ and 0 are zero matrices with proper dimensions. Let $\boldsymbol{X} = \begin{bmatrix} -\boldsymbol{F} \\ \boldsymbol{I}_{n-r} \end{bmatrix}$. Then we have

$$\boldsymbol{R}\boldsymbol{X} = \begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{F} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -\boldsymbol{F} \\ \boldsymbol{I}_{n-r} \end{bmatrix} = 0$$

Therefore, each column $\boldsymbol{X}$ is a special solution to the system $\boldsymbol{A}x = 0$. Next, we want to show other solutions to $\boldsymbol{A}x = 0$ are linear combinations of those special solutions.
Suppose there is a $x = (x_1, x_2) \in N(\boldsymbol{A})$. Then

$$\boldsymbol{R}x = \begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{F} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + \boldsymbol{F} x_2 \\ 0 \end{bmatrix} = 0.$$

That implies $x_1 = -\boldsymbol{F} x_2$ and $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\boldsymbol{F} \\ \boldsymbol{I}_{n-r} \end{bmatrix} x_2 = \boldsymbol{X} x_2$. Therefore, arbitrary $x \in N(\boldsymbol{A})$ is a linear combination of special solutions, i.e. $C(\boldsymbol{X}) = N(\boldsymbol{A})$.
Since $\boldsymbol{X}$ contains an identity matrix, the special solutions are independent and the rank of $C(\boldsymbol{X})$ is $n - r$. Therefore, the rank of $N(\boldsymbol{A})$ is $n - r$. □

Recall that the columns in $\boldsymbol{A}$ correspond to the $\boldsymbol{I}_r$ in $\boldsymbol{R}$ are called pivot columns and other columns are free columns. In $\boldsymbol{A}x = b$, the variables in x that correspond to pivot columns are called *pivot variables* and others are *free variables*. As the proof shows, the way to find special solutions of $\boldsymbol{A}x = 0$ is to let one of the free variables is 1 while other free variables are 0 and then solve the equation $\boldsymbol{A}x = 0$. There are total $n - r$ special solutions.

2.2 Solving $Ax = b$: complete solutions

Lemma 5 (Solvability of $\boldsymbol{A}x = b$)**.** *The system $\boldsymbol{A}x = b$ is solvable only when $b \in C(\boldsymbol{A})$.*

Proof. For any x, $\boldsymbol{A}x \in C(\boldsymbol{A})$. Therefore, to satisfy $\boldsymbol{A}x = b$, $b \in C(\boldsymbol{A})$. $\square$

Lemma 6 (Complete solution)**.** *The complete solution of $\boldsymbol{A}x = b$ is given by $x_{comp} = x_p + x_n$, where x_p is a particular solution that $\boldsymbol{A}x_p = b$ and $x_n \in N(\boldsymbol{A})$.*

Proof. Suppose $x = x_p + x_0$ is an arbitrary solution to $\boldsymbol{A}x = b$. Then we have

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

Then $x_0 \in N(\boldsymbol{A})$. $\square$

One way to find a particular solution is to let all the free variables be 0 and solve the equations.

Here is a summary table that discusses the rank of $\boldsymbol{A}$, the form of $\boldsymbol{R}$ and the situation about the solutions.

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