Linear Algebra

A summary for MIT 18.06SC

Jiaxin Hu

June 29, 2020

1 Matrices & Spaces

1.1 Basic concepts

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

Lemma 1 (Communicative matrix multiplication). For matrices $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times n}$, the matrix multiplication of A and B is communicative, i.e. AB = BA, if all the normalized eigenvectors of A and B are identical.

Proof. If all the normalized eigenvectors of A and B are identical, there exists a matrix $Q \in \mathbb{R}^{n \times n}$ such that

$$A = QD_AQ^{-1}, \quad B = QD_BQ^{-1},$$

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

$$AB = QD_AQ^{-1}QD_BQ^{-1} = QD_AD_BQ^{-1} = QD_BD_AQ^{-1} = QD_BQ^{-1}QD_AQ^{-1} = BA.$$

Therefore, the matrix multiplication of A and B is communicative.

1.2 Permutation of matrices

For any matrix A, we can swap its rows by multiplying a permutation matrix P on the left of A. For example,

$$m{PA} = egin{bmatrix} 0 & 0 & 1 \ 1 & 0 & 0 \ 0 & 1 & 0 \end{bmatrix} egin{bmatrix} a_1 \ a_2 \ a_3 \end{bmatrix} = egin{bmatrix} a_3 \ a_1 \ a_2 \end{bmatrix}$$

where a_k refers to the k-th row of A. The inverse of permutation matrix P is $P^{-1} = P^T$, 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 can also swap the columns of the matrix A by multiplying a permutation matrix on the right of 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 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 U is an upper traingular matrix.

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

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

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 next column.

We write $E_{32}(E_{21}A) = (E_{32}E_{21})A = U$ because matrix multiplication is associative. Let 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 A. Therefore, the elimination process of A is

$$EPA = U, (1)$$

where U is an upper triangular matrix.

Next, we prove the invertibility of the elimination matrix.

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

Proof. We write the elimination matrix as

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

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

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

Therefore, $I_n - ce_i e_i^T$ is the inverse of E_{ij} . The elimination matrix E_{ij} is invertible.

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

Proof. By the proof of lemma 2, the matrix E_{ij} and its inverse are written as

$$\boldsymbol{E}_{ij} = \boldsymbol{I}_n + ce_i e_j^T, \quad \boldsymbol{E}_{ij}^{-1} = \boldsymbol{I}_n - ce_i e_j^T.$$

WLOG, we assume E_{ij} is a lower-triangular matrix. Then $ce_ie_j^T$ and $-ce_ie_j^T$ are also lower-triangular matrices. Therefore, E_{ij}^{-1} is a lower-triangular matrix.

1.4 Gauss-Jordan Elimination

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

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

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n. \tag{2}$$

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

We call this elimination process of finding **E** as Gauss-Jordan Elimination.

1.5 Factorization of matrices

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

$$PA = E^{-1}U$$

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

$$PA = LU, (3)$$

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

1.6 Time complexity of elimination

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 repeat single eliminations (n-1) times, and thus require $\mathcal{O}(n^2)$ operations. Similarly, we require $\mathcal{O}((n-1)^2)$ to eliminate the elements below the second diagonal element. Repeat the elimination until we meet the *n*-th diagonal element. Therefore, we require $\mathcal{O}(n^3)$ operations to obtain an upper-triangular matrix by elimination:

$$1^{2} + 2^{2} + \dots + (n)^{2} = \sum_{i=1}^{n} i^{2} \approx \int_{0}^{n} x^{2} dx = \frac{1}{3} n^{3} = \mathcal{O}(n^{3}).$$

1.7 Reduced row echelon form of matrices

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

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

There is another example,

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

Assume 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*.

1.8 Vector space, Subspace and Column space

- Vector space is a collection of vectors that is 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 a set that contains only a zero vector.

- Vectors $v_1, ..., v_n$ span a space that consists all the combination of those vectors.
- Column space of matrix A is the space spanned by the columns of A. Let C(A) denote the column space of A.

If $v_1, ..., v_n$ span a space S, then S is the smallest space that contain those vectors.

- Basis of a vector space is a sequence of vectors $v_1, ..., v_n$ that satisfy: (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. Let dim(A) denote the dimension of space A.

1.9 Matrix rank

The rank of a matrix \boldsymbol{A} is defined as the dimension of the columns space of \boldsymbol{A} . Rank of matrix \boldsymbol{A} is also equal to the number of pivot columns of \boldsymbol{A} . Let $rank(\boldsymbol{A})$ denote the rank of matrix \boldsymbol{A} . We have

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

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

The rank of a square matrix is closely related to its 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. First, we assume A is full rank and prove the invertibility of A.

Consider the RREF form of \mathbf{A} , $\mathbf{R} = RREF(\mathbf{A})$. There are elimination matrix \mathbf{E} and permutation matrix \mathbf{P} such that

$$EPA = R$$
.

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

Second, we assume A is invertible and prove the full rankness of A by contradiction.

Assume \mathbf{A} has an inverse \mathbf{A}^{-1} and $rank(\mathbf{A}) < n$. By equation (4), $dim(C(\mathbf{A})) = rank(\mathbf{A}) < n$, which implies that the columns of \mathbf{A} are linearly dependent. Then, there exists a non-zero vector v such that

$$\mathbf{A}v = 0. \tag{5}$$

Multiplying A^{-1} on both sides of equation (5), we have

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

However, it contradicts the fact that v is non-zero. Therefore, A is full rank.

The rank of A also effects the number of solutions to the system Ax = b. We will discuss it in next section.

2 Solving Ax = b

Here we discuss the solutions of the linear system 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 $\mathbf{A}x = 0$. Let $N(\mathbf{A})$ denote the nullspace of \mathbf{A} .

Lemma 4 (Nullspace). The nullspace of matrix **A** is a vector space.

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

$$\mathbf{A}(c_1v_1 + c_2v_2) = c_1\mathbf{A}v_1 + c_2\mathbf{A}v_2 = 0, \quad \forall c_1, c_2 \in \mathbb{R}.$$
 (6)

The equation (6) implies that any linear combination of vectors in $N(\mathbf{A})$ is also a vector in $N(\mathbf{A})$. Therefore, $N(\mathbf{A})$ is closed to linear combination.

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

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 arbitrary matrix. Let $\mathbf{X} = \begin{bmatrix} -\mathbf{F} \\ \mathbf{I}_{n-r} \end{bmatrix}$. 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 of X is a special solution to the system Ax = 0. Next, we show other solutions to Ax = 0 are linear combinations of those special solutions. Suppose there is a solution $x = (x_1, x_2) \in N(A)$, where $x_1 \in \mathbb{R}^r$ and $x_2 \in \mathbb{R}^{n-r}$. We have

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

That implies $x_1 = -\mathbf{F}x_2$ 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 special solutions, i.e. $C(\mathbf{X}) = N(\mathbf{A})$. Therefore, the rank of nullspace is $rank(N(\mathbf{A})) = dim(C(\mathbf{X})) = n - r$.

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

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 those special solutions by assigning 1 to a free variable and 0 to other free variables, and then solving the system $\mathbf{A}x = 0$.

2.2 Solving Ax = b: complete solutions

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

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

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

Proof. Suppose $x = x_p + x_0$ is an arbitrary 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,
$$x_0 \in N(\mathbf{A})$$
.

Usually, we find a particular solution by assigning 0 to free variables, and solving the system $\mathbf{A}x = b$. The following table discusses the rank of \mathbf{A} , the form of \mathbf{R} , dimension of nullspace $N(\mathbf{A})$, and solutions of $\mathbf{A}x = b$.

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