Linear independence and systems of linear equations

Francesco Preta
07/21/2020

The coefficient matrix as a linear operator

Consider the general system of linear equations $A\mathbf{x} = \mathbf{b}$. Finding a solution for the system corresponds to finding a vector $\mathbf{x} \in \mathbb{R}^n$ which is mapped to a given vector $\mathbf{b} \in \mathbb{R}^m$. In this sense, there are three possibilities:

1. $\mathbf{b} \in$ Im(A), Ker$(A) = \{0\}$, in which case there exists a left inverse A^{-1} granting a unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.
2. $\mathbf{b} \in$ Im(A), Ker$(A) \neq \{0\}$, in which case there are infinitely many solutions.
3. $\mathbf{b} \notin$ Im(A), in which case there are no solutions.

Linear independence

Definition

Let $\mathscr{V}$ be a vector space and $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ a subset of vectors in $\mathscr{V}$.
We say that $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$are **linearly independent** if, for

$$\alpha_1, ..., \alpha_k \in \mathbb{R}$$

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + ... + \alpha_k \mathbf{v}_k = 0 \implies (\alpha_1, ..., \alpha_k) = \mathbf{0}$$

In other word, $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ are linearly independent if any nonzero linear combination of them is also nonzero.

Consequences of the definition

- No single vector $\mathbf{v}_j \in \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a linear combination of the other $k-1$ vectors.
- Two vectors are linearly independent if one is not a multiple of the other.
- If $\mathbf{0}$ is in $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$, then the vectors are not independent, since any linear combination attributing a nonzero weight only to $\mathbf{0}$ would give $\mathbf{0}$ as a result.

How to determine linear independence?

For a given set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ of $\mathbb{R}^n$, we need to be able to determine whether they're independent or not. This corresponds to asking if we can find $(\alpha_1, \ldots, \alpha_k) \neq \mathbf{0}$such that

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_k\mathbf{v}_k = 0$$

This condition corresponds to solving the homogeneous system $V\mathbf{a} = \mathbf{0}$ for

$$V = [\mathbf{v}_1 \qquad \mathbf{v}_2 \qquad \cdots \qquad \mathbf{v}_k] \qquad \mathbf{a} = \begin{bmatrix} \alpha_1 \\ \cdots \\ \alpha_k \end{bmatrix}$$

Example

Determine if the following vectors are linearly independent.

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \\ -1 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 2 \\ -2 \\ 0 \end{bmatrix}, \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -2 \\ 8 \\ -3 \end{bmatrix}.$$

Example:

Determine if the following vectors are linearly independent.

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \\ -1 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 2 \\ -2 \\ 0 \end{bmatrix}, \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ -3 \end{bmatrix}.$$

Basis and Dimension

Definition

Let $\mathscr{V}$ be a vector space. A set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is called a **basis** for $\mathscr{V}$ if

- $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a set of linearly independent vector.
- For every $\mathbf{v} \in \mathscr{V}$ there exists $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ such that

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n$$

In other words, $\mathscr{V} = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$

The **dimension** of $\mathscr{V}$ is the cardinality of its basis.

Example: Canonical basis

For example, for $\mathscr{V} = \mathbb{R}^n$, the canonical basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ constitute a basis. The dimension is n, as expected.

Maximal set of linearly independent vectors

Theorem

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a subset of a n-dimensional vector space. Then if $k > n$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is not linearly independent.

In fact, there are at most n linearly independent vectors in a n-dimensional vector space and any group of them constitutes a basis.

Vector subspace

Definition

Let $\mathscr{V}$ be a finite dimensional vector space. A subset $\mathscr{V}' \subset \mathscr{V}$ is called a vector subspace of $\mathscr{V}$ if the following conditions are true:

- $\mathbf{0} \in \mathscr{V}'$.
- For every $\mathbf{v}_1, \mathbf{v}_2 \in \mathscr{V}'$ and $\alpha_1, \alpha_2 \in \mathbb{R}$, $\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 \in \mathscr{V}'$.

In other words, a vector subspace is a subset of a vector space which is closed under linear combinations.

Subspaces of $\mathbb{R}^n$

Given a set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subset \mathbb{R}^n$, span$(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ is a vector subspace of $\mathbb{R}^n$. If $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ are linearly independent, then span$(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ is a k-dimensional vector subspace of $\mathbb{R}^n$. Otherwise, span$(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ has dimension $d < k$.

Linear operators

Theorem

Let A be a $m \times n$ matrix. Then $Ker(A)$ is a vector subspace of $\mathbb{R}^n$ and Im(A) is a vector subspace of $\mathbb{R}^m$ spanned by the columns of A.

Proof:

Finding a basis if you know a set of generators

Theorem

Let $\mathscr{V}$ be a vector space, $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subset \mathscr{V}$ and $H = \text{span}(v_1, \ldots, v_k)$. Then

- If $\mathbf{v}_j$ is a linear combination of the other $\mathbf{v}_i$'s, then $H = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_k\}$.
- If $\mathbf{v}$ is a non-trivial linear combination of $\mathbf{v}_l$ with $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$, then $H = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{l-1}, \mathbf{v}, \mathbf{v}_{l+1}, \ldots, \mathbf{v}_k\}$.
- If $H \neq \{\mathbf{0}\}$, a subset of $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a basis for H.

Finding a basis if you know a set of generators

In order to find a basis for a vector space H, we need to remove $\mathbf{0}$ and linear combinations from the vectors that span H.

Example:

Find a basis for Ker(A) and Im(A) for

$$A = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 4 & 2 & -1 & 0 \\ 2 & 2 & -1 & 2 \end{bmatrix}$$

Span of pivot columns

Theorem

$\text{Im}(A)$ is spanned by the pivot columns of A.

This happens because the pivot columns correspond to non-free variables whose only feasible solution for the homogeneous system $A\mathbf{x} = \mathbf{0}$ is given by $\mathbf{0}$, once the free variables are eliminated.

Rank of matrices

Rank of a matrix

Definition

Let A be a $m \times n$ matrix. Then rank(A) is the dimension of the linear subspace of $\mathbb{R}^m$ generated by the column vectors $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$.

Equivalent definitions

Theorem

Let A be a $m \times n$ matrix. Then the following are equivalent:

- rank$(A) = r$.
- The subspace of $\mathbb{R}^n$ generated by the row vectors of A has dimension r.
- There exists a square $r \times r$ submatrix of A with nonzero determinant and every square submatrix of higher dimension has determinant equal to zero.

Example:

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 2 & -2 \\ 2 & -2 & 8 \\ -1 & 0 & -3 \end{bmatrix}$$

Simplifying calculations:

Theorem

Let A be a $n \times m$ matrix and A' be a r-dimensional square submatrix with nonzero determinant. Then rank$(A) = r$ if for every row I and column j not already in A', the matrix obtained by adding such row and column to A' has zero determinant.

Linear independence and basis through rank

In order to understand whether a few vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ are linearly independent we can do the following:

1. Consider the matrix V having $\mathbf{v}_i$ as i-th column.
2. Calculate the rank of V through determinants of square submatrices.
3. A basis for span$(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ is given by the columns included in the square submatrix of maximum rank.

Rank-kernel Theorem

Let A be a $m \times n$ matrix. Then rank$(A) + \dim($Ker$(A)) = n$

Proof

Consider the homogeneous system $A\mathbf{x} = \mathbf{0}$. The general form of its solution in $\mathbb{R}^n$ can be divided into free variables and non-free variables. The number of free variables correspond to the dimension of Ker(A), while the number of non-free variables correspond to the number of pivot columns. Since this last value is captured by the rank, the theorem is proved.

Previous example:

$$A = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 4 & 2 & -1 & 0 \\ 2 & 2 & -1 & 2 \end{bmatrix}$$

Properties of the rank:

- rank$(A) =$ rank(A^T).
- rank$(A) \leq \min\{m, n\}$.
- If B is obtained by row-reduction operations on A, then rank$(B) =$ rank(A).
- If B is obtained by column-reduction operations on A, then rank$(B) =$ rank(A).
- If $m = n$ then rank$(A) = n$ if and only if A is invertible.

Revisiting last week's results

A square system $A\mathbf{x} = \mathbf{b}$ with nonzero determinant always has a unique solution since $\det(A) \neq 0$ is equivalent to rank$(A) = n$ and thus, by the Rank-Kernel theorem Ker$(A) = \{\mathbf{0}\}$. In this way, A is bijective and for any $\mathbf{b} \in \mathbb{R}^n$ there exists a unique linear combination of the columns of A with coefficients $x_1, \ldots, x_n$ such that $\mathbf{b} = x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n$.

Linear independence in systems of linear equations

Rouché-Capelli Theorem

Consider a system of linear equation having matrix form $A\mathbf{x} = \mathbf{b}$. Then the system admits solutions if and only if the coefficient matrix A and the augmented matrix $A' = [A \,|\, \mathbf{b}]$ have the same rank. If that is the case, the number of free variables is given by $n - \text{rank}(A)$, where n is the total number of variables.

Proof:

Underdetermined systems $(m < n)$

In case $\text{rank}(A) = m$, adding the column matrix $\mathbf{b}$ will not have an impact on the system, since

$m \leq \text{rank}([A \,|\, \mathbf{b}) \leq \min\{m, n+1\} = m$.

By Rouché-Capelli's Theorem, the system will admit solutions. In particular, $n - m$ variables will be free, while the rest will be determined.

Undetermined systems (continued)

In case rank$(A) = r < m$, then the existence of solution depends on rank$([A \,|\, \mathbf{b}])$, which could potentially be larger than r if $\mathbf{b}$ is not a linear combination of the columns of A. In general, if solutions exist, $n - r$ variables will be free, while the rest will be determined.

Example:

$$\begin{bmatrix} 2 & -1 & 0 & 1 \\ -1 & 3 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Example:

$$\begin{bmatrix} 2 & -1 & 0 & 1 \\ -1 & 3 & 1 & 0 \\ 0 & 5 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Square systems $(m = n)$

In this case, being maximum rank coincides with having nonzero determinant, which is exactly the case we dealt with in the last lecture. Notice that in this case

$$n = \text{rank}([A \mid \mathbf{b}]) \leq \min n, n + 1 = n$$

so that the condition of Rouché-Capelli's Theorem is respected.

Square systems (continued)

In case rank$(A) = r < n$, we have two possibilities depending on rank$([A \,|\, \mathbf{b}])$. If rank$([A \,|\, \mathbf{b}]) = r$, then the system has infinitely many solutions and $n - r$ variables will be free. Otherwise, the system will have no solution.

Example:

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 1 \\ -2 & -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Overdetermined systems $(m > n)$

In case $m > n$, the maximum rank of the coefficient matrix is n. Even in the maximum rank case, adding the column $\mathbf{b}$ could lead to an increase in the rank of the augmented matrix. This happens because there are more equations than unknowns. Ideally, every additional equation should be a linear combination of all the previous one, so that we don't add stringent requirements that cannot be satisfied by the unknowns.

Example:

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 3 \\ -1 & 2 & 1 \\ 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 1 \end{bmatrix}$$

Inverses for rectangular matrices

Left and right inverses

Theorem

Let A be a $m \times n$ matrix. Then

- There exists a right inverse Y such that $AY = I$ if and only if the rows of A are linearly independent. This is generally not unique.
- There exists a left inverse X such that $XA = I$ if and only if the columns of A are linearly independent. This is generally not unique.

Left and right inverse (continued)

Notice that for wide matrices ($n > m$), columns cannot be linearly independent, therefore they only admit right inverses, while in tall matrices ($n < m$) rows cannot be linearly independent and so they only admit left inverses. In any case, a matrix A admits a (left and/or right) inverse if and only if it has maximum rank.

Applications to under-determined systems

If the system is under-determined ($m > n$) and A has maximum rank, then A admits a right inverse Y. Since A has maximum rank, by Rouché-Capelli's Theorem, the system admits infinitely many solutions. Then, let $\mathbf{x} = Y\mathbf{b}$, we have$A\mathbf{x} = AY\mathbf{b} = \mathbf{b}$, therefore $\mathbf{x}$ is a solution to the system. Since Y is not unique, we can find other solutions by finding different right inverses.

Applications to overdetermined systems

If the system is over-determined ($n > m$) and A has maximum rank, there is no guarantee that the system has solutions. However, if it does, it has only one solution. If $\mathbf{x}$ is a solution of the system, then $XA\mathbf{x} = X\mathbf{b}$, meaning that $\mathbf{x} = X\mathbf{b}$ is the unique desired solution. On the other hand, if $AX\mathbf{b} \neq \mathbf{b}$, then there are no solutions.

Pseudo-inverses

Theorem

A admits a left inverse if and only if the square matrix $A^T A$ is invertible. In this case, a left inverse for A is given by $A^\dagger = (A^T A)^{-1} A^T$

A admits a right inverse if and only if the square matrix AA^T is invertible. In this case, a right inverse for A is given by $A^\dagger = A^T (AA^T)^{-1}$

The matrices above are called **pseudo-inverses** of A.

Example:

Solve $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 1 & -2 \\ 3 & 1 \\ 1 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} \text{ or } \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$