Lecture 18: The Rank of a Matrix and Consistency of Linear Systems

Winfried Just
Department of Mathematics, Ohio University

February 28, 2018

Review: The linear span

Definition

Let $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_n\}$ be a set of vectors of the same order. The linear span of these vectors is the set $span(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_n)$ of all linear combinations of the vectors $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_n$.

Remark on terminology: Some textbooks use the terminology "$\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_n\}$ is a spanning set for a set of vectors V."

This means the same thing as "$span(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_n) = V$."

Some textbooks also write $S(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_n)$ instead of $span(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_n)$.

Review: How to recognize linear (in)dependence?

Let $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_n\}$ be a set of column vectors of the same dimension.

We are looking for coefficients $c_1, c_2, \ldots, c_n$ such that

$c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + \cdots + c_n\vec{\mathbf{v}}_n = \vec{\mathbf{0}}$.

This boils down to solving a system of linear equations

$\mathbf{A}\vec{\mathbf{c}} = \vec{\mathbf{0}}$,

where the columns of $\mathbf{A}$ are $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_n$,

and $\vec{\mathbf{c}} = [c_1 \ c_2 \ \ldots \ c_n]^T$.

If the system is underdetermined the vectors $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_n$ are linearly dependent; otherwise they are linearly independent.

An example

Are the following vectors linearly independent?

$$\vec{\mathbf{v}}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \vec{\mathbf{v}}_2 = \begin{bmatrix} -2 \\ -4 \\ -6 \end{bmatrix} \qquad \vec{\mathbf{v}}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

To find all c_1, c_2, c_3 with $c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3 = \vec{\mathbf{0}}$,

we need to solve the following system of linear equations:

$$\begin{aligned} c_1 - 2c_2 + c_3 &= 0 \\ 2c_1 - 4c_2 &= 0 \\ 3c_1 - 6c_2 + c_3 &= 0 \end{aligned}$$

We perform Gaussian elimination on the augmented matrix

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 2 & -4 & 0 & 0 \\ 3 & -6 & 1 & 0 \end{bmatrix} \underrightarrow{\text{subtract 2(row 1) from row 2}} \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 3 & -6 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 3 & -6 & 1 & 0 \end{bmatrix} \underrightarrow{\text{subtract 3(row 1) from row 3}} \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & -2 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & -2 & 0 \end{bmatrix} \underrightarrow{\text{subtract row 2 from row 3}} \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \underrightarrow{\text{divide row 2 by -2}} \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Finding all solutions

The row-reduced matrix

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

is the augmented matrix of the equivalent system

$$\begin{aligned} c_1 - 2c_2 + c_3 &= 0 \\ c_3 &= 0 \\ 0 &= 0 \end{aligned}$$

Every vector $[2c_2, c_2, 0]^T$ is a solution.

The vectors $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3$ are linearly dependent.

What can happen in principle?

Let $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3$ be 3×1 be (nonzero) column vectors. To find all $\vec{\mathbf{c}} = [c_1, c_2, c_3]^T$ with $c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3 = \vec{\mathbf{0}}$, we perform Gaussian elimination. The row-reduced matrix may then look like:

$\begin{bmatrix} 1 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ Underdetermined. $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3$ are linearly dependent.

$\begin{bmatrix} 1 & ? & ? & 0 \\ 0 & 1 & ? & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ Underdetermined. $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3$ are linearly dependent.

$\begin{bmatrix} 1 & ? & ? & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ Underdetermined. $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3$ are linearly dependent.

$\begin{bmatrix} 1 & ? & ? & 0 \\ 0 & 1 & ? & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ $\vec{\mathbf{c}}$ must be $\vec{\mathbf{0}}$. $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3$ are linearly independent.

Row rank and column rank of a matrix

Definition

The row rank of a matrix is the maximum size of a linearly independent subset of its row vectors.
The column rank of a matrix is the maximum size of a linearly independent subset of its column vectors.

Example: Both the row rank and the column rank of a zero matrix $\mathbf{O}$ are 0.

Example: Let $\mathbf{A}$ be an $m \times n$ matrix.

The row vectors of $\mathbf{A}$ form a linearly independent set if, and only if, the row rank of $\mathbf{A}$ is m.

The column vectors of $\mathbf{A}$ form a linearly independent set if, and only if, the column rank of $\mathbf{A}$ is n.

More examples: Our row-reduced matrices

Consider the row-reduced matrices that we could obtain when testing linear independence of a set $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3\}$ of 3×1 (nonzero) column vectors.

$$\begin{bmatrix} 1 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$ Row rank = column rank = 1.

$$\begin{bmatrix} 1 & ? & ? & 0 \\ 0 & 1 & ? & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$ Row rank = column rank = 2.

$$\begin{bmatrix} 1 & ? & ? & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$ Row rank = column rank = 2.

$$\begin{bmatrix} 1 & ? & ? & 0 \\ 0 & 1 & ? & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$ Row rank = column rank = 3.

Homework 48: Verify the claim about row rank and column rank made on the previous slide for the third matrix by considering possible coefficients that will give $\vec{\mathbf{0}}$ as a linear combination of some row or column vectors.

Theorem

If $\mathbf{B}$ *is obtained from* $\mathbf{A}$ *by an elementary row operation, then the row rank of* $\mathbf{B}$ *is equal to the row rank of* $\mathbf{A}$ *and the column rank of* $\mathbf{B}$ *is equal to the column rank of* $\mathbf{A}$*.*

Corollary

The process of Gaussian elimination preserves both row rank and column rank.

The rank of a matrix

The examples on slide 9 illustrate the following result:

Theorem

Both the row rank and the column rank of a row-reduced matrix are equal to its number of nonzero rows.

Corollary

The row rank and the column rank of every matrix $\mathbf{A}$ *are equal.*

We call this common number the rank of $\mathbf{A}$. It is denoted by $r(\mathbf{A})$.

An $n \times n$ square matrix $\mathbf{A}$ is said to have full rank if $r(\mathbf{A}) = n$, that is, if its column vectors (equivalently: its row vectors) form an independent set.

Examples of full rank submatrices

Submatrices of full rank for the following row-reduced matrices are shown in red. Depending on the values of the unspecified entries there may be more such submatrices.

$$\begin{bmatrix} \color{red}{1} & ? & ? & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{Rank} = 1.$$

$$\begin{bmatrix} \color{red}{1} & \color{red}{?} & ? & 0 \\ \color{red}{0} & \color{red}{1} & ? & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{Rank} = 2.$$

$$\begin{bmatrix} \color{red}{1} & ? & \color{red}{?} & 0 \\ \color{red}{0} & 0 & \color{red}{1} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{Rank} = 2.$$

$$\begin{bmatrix} \color{red}{1} & \color{red}{?} & \color{red}{?} & 0 \\ \color{red}{0} & \color{red}{1} & \color{red}{?} & 0 \\ \color{red}{0} & \color{red}{0} & \color{red}{1} & 0 \end{bmatrix} \qquad \text{Rank} = 3.$$

A characterization of the rank of a matrix

Theorem

The rank $r(\mathbf{A})$ of a matrix is the largest n such that $\mathbf{A}$ contains a submatrix $\mathbf{B}$ of order $n \times n$ and full rank.

Homework 49: Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$

Suppose Gaussian elimination is performed by switching rows 1 and 3 in the first step and only using the other two elementary row operations in later steps. If the resulting row-reduced form is

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

find a submatrix $\mathbf{B}$ of $\mathbf{A}$ of full rank.

Homework 50: For each of the following matrices, find its rank. Which of these matrices have full rank?

(a) $\mathbf{A} = \begin{bmatrix} -1 & 2 \\ -2 & 3 \end{bmatrix}$

(b) $\mathbf{B} = \begin{bmatrix} 3 & 2 & 1 \\ 4 & 5 & 6 \\ 1 & 1 & 1 \end{bmatrix}$

(c) $\mathbf{C} = \begin{bmatrix} 3 & 2 & 1 & 0 \\ 1 & 5 & 0 & 6 \\ -1 & 1 & -1 & 1 \\ 5 & 6 & 2 & 5 \end{bmatrix}$

An (informal) geometric interpretation of the rank

Theorem

Suppose $\mathbf{A}$ *is an* $m \times n$ *matrix. Then the linear span of the row vectors of* $\mathbf{A}$ *is a subspace of geometric dimension* $r(\mathbf{A})$ *of* $\mathbb{R}^n$ *and the linear span of the column vectors of* $\mathbf{A}$ *is a subspace of geometric dimension* $r(\mathbf{A})$ *of* $\mathbb{R}^m$.

Example: Let $\mathbf{A} = \begin{bmatrix} 1 & 2 & -3 \\ 4 & 5 & -9 \\ 0 & 0 & 0 \end{bmatrix}$

Then $span([1,4,0]^T, [2,5,0]^T, [-3,-9,0]^T)$ is the x-y-plane,

and $span([1,2,-3],[4,5,-9],[0,0,0])$ is the plane of all vectors

$[x,y,z]$ such hat $x+y+z=0$.

Homework 51: Verify the last two assertions.

Review: Consistency of systems of linear equations

Let $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ be the matrix form of a system of linear equations.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

The following statements are equivalent, that is, express the same property in different ways:

- The system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is consistent.
- $\vec{\mathbf{b}} = x_1\vec{\mathbf{a}}_1 + x_2\vec{\mathbf{a}}_2 + \cdots + x_n\vec{\mathbf{a}}_n$ is a linear combination of the column vectors $\vec{\mathbf{a}}_1, \vec{\mathbf{a}}_2, \ldots, \vec{\mathbf{a}}_n$ of $\mathbf{A}$.
- $\vec{\mathbf{b}}$ is in $span(\vec{\mathbf{a}}_1, \vec{\mathbf{a}}_2, \ldots, \vec{\mathbf{a}}_n)$.

Consistency and the rank

Let $\mathbf{A}\vec{x} = \vec{\mathbf{b}}$ be the matrix form of a system of linear equations, and let $[\mathbf{A}, \vec{\mathbf{b}}]$ be its augmented matrix.

Theorem

The system $\mathbf{A}\vec{x} = \vec{\mathbf{b}}$ is consistent if, and only if, $r(\mathbf{A}) = r([\mathbf{A}, \vec{\mathbf{b}}])$.

Idea of the proof: Consistency requires that $\vec{\mathbf{b}}$ is in the linear span of the column vectors of $\mathbf{A}$. Thus adding the last column $\vec{\mathbf{b}}$ to the coefficient matrix $\mathbf{A}$ cannot add a new dimension to the linear span and cannot increase the rank of the matrix.

Corollary

When $r(\mathbf{A}) = m$ is equal to the number of rows of $\mathbf{A}$, then every system of the form $\mathbf{A}\vec{x} = \vec{\mathbf{b}}$ is consistent.

Proof: Since $[\mathbf{A}, \vec{\mathbf{b}}]$ has also m rows, $r([\mathbf{A}, \vec{\mathbf{b}}]) \leq m$, and $r(\mathbf{A}) = m$ implies that $r(\mathbf{A}) = r([\mathbf{A}, \vec{\mathbf{b}}])$.

An application to chemical reaction systems

Suppose $\mathbf{S}$ is the stoichiometric matrix of a chemical reaction system. We have seen that if we write the reaction vectors $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n$ as columns of $\mathbf{S}$, then the vector $\vec{\mathbf{w}}$ of net changes is given as $\mathbf{S}\vec{\mathbf{k}}$, where $\vec{\mathbf{k}} = [k_1, \dots, k_n]^T$ is the vector of net rates at which the forward reactions occur.

Thus $\vec{\mathbf{w}} = k_1\vec{\mathbf{v}}_1 + \cdots + k_n\vec{\mathbf{v}}_n$

is a linear combination of the reaction vectors and must be in the linear span $span(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n)$.
This linear span of the set of reaction vectors is called the stoichiometric subspace.

Homework 52: Suppose you have a stoichiometric matrix of order $m \times n$ that represents n reactions between m chemical species in a closed reaction system (without net inflow, net outflow, or contributions from or to other reactions). Prove that $r(\mathbf{S}) < m$.

Revisiting our chemical reaction system

Recall our chemical reaction system

$A + 2B \rightleftarrows 2C \qquad A + 2C \rightleftarrows 2D$

$A + B \rightleftarrows D \qquad B + D \rightleftarrows 2C$

If we write reaction vectors as columns, then he stoichiometric matrix can be written as

$$\mathbf{S} = \begin{bmatrix} -1 & -1 & -1 & 0 \\ -2 & 0 & -1 & -1 \\ 2 & -2 & 0 & 2 \\ 0 & 2 & 1 & -1 \end{bmatrix}$$

Homework 53: (a) Find the rank of the stoichiometric matrix on the previous slide.

(b) For each of the following vectors, determine whether it could possibly be the vector of net concentration changes for the reaction system on the previous slide.

$$\vec{\mathbf{w}}_1 = \begin{bmatrix} -3 \\ -2 \\ -2 \\ 4 \end{bmatrix} \qquad \vec{\mathbf{w}}_2 = \begin{bmatrix} -2 \\ -2 \\ 1 \\ 3 \end{bmatrix}$$

(c) If the vector of net changes of point (b) that is feasible has been observed, can we conclude that at least 2 of the reactions did occur? Can we conclude that at least 3 of the reactions did occur? Or even that all 4 reactions did occur?