Conversation 29: The Rank and Theory of Solutions

Winfried Just
Department of Mathematics, Ohio University

MATH3200: Applied Linear Algebra

What is the rank of a matrix good for?

Denny: Why do we have to learn this rank of a matrix stuff?

Theo: It is an important concept in linear algebra.

Bob: And it is easy to compute, as we practiced in Module 49.

Frank: Yeah, but what can we do with it? This is a course for engineering students, remember?

Alice: The rank gives you important information about how you can represent the solution set of a linear system.

Cindy: You mean, like, writing down the solution set when there are more than one solutions? I always find that difficult.

Alice: Yes, the rank is especially useful in this case.

Frank: How is that?

Alice: Consider a system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ of linear equations

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

Let $\mathbf{A}$ and $[\mathbf{A}, \vec{\mathbf{b}}]$ be the coefficient matrix and the augmented matrix of this system, respectively. In Lecture 27 we learned the following fact:

Theorem

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

Does the system have any solutions?

Theo: It now follows immediately that by comparing $r(\mathbf{A})$ and $r([\mathbf{A}, \vec{\mathbf{b}}])$, we can determine whether the solution set is empty, or nonempty.

Denny: Give us a break with your highfalutin terminology, Theo!

Theo: "Empty" means that the solution set has 0 elements, so that the system is inconsistent; "nonempty" means that there is at least one solution, so that the system is consistent.

Frank: Big deal! But in order to compare $r(\mathbf{A})$ with $r([\mathbf{A}, \vec{\mathbf{b}}])$, I would need to compute the ranks of two matrices by doing Gaussian elimination on both.

Seems much more work than solving the system in the way we learned in Chapter 2, by doing just one Gaussian elimination.

Can we get away with computing $r(\mathbf{A})$ only?

Denny: But perhaps we could get away with just finding the rank of the coefficient matrix $\mathbf{A}$? That would be even less work, especially if we were interested in the consistency of systems $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ for the same $\mathbf{A}$ and a lot of different vectors $\vec{\mathbf{b}}$. Would that work, Theo?

Theo: Sometimes it would, sometimes not.

Denny: So when would it work?

Theo: When the coefficient matrix $\mathbf{A}$ is of order $m \times n$ and $r(\mathbf{A}) = m$, the number of rows of $\mathbf{A}$.

Then we can use the corollary from the end of Lecture 27:

Corollary

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

When can we get away with computing $r(\mathbf{A})$ only?

Denny: Ah, great! Means we only need to compute $r(\mathbf{A})$. When $r(\mathbf{A})$ is equal to the number of rows of $\mathbf{A}$, the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is consistent, and when $r(\mathbf{A})$ is less than the number of rows of $\mathbf{A}$, the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is inconsistent!

Question C29.1: What do you think of Denny's observation?

Bob: For a given coefficient matrix $\mathbf{A}$, there are many possible vectors $\vec{\mathbf{b}}$. Which one do you mean when you talk about *the* system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$?

Denny: Any $\vec{\mathbf{b}}$. I mean: Look at any system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ where $\mathbf{A}$ has order $m \times n$ and $\vec{\mathbf{b}}$ is a column vector of length m. When $r(\mathbf{A}) = m$, consistent. When $r(\mathbf{A}) < m$, the system is inconsistent.

Cindy: But when $\vec{\mathbf{b}} = \vec{\mathbf{0}}$ is the $m \times 1$ zero vector, doesn't the system then always have a solution?

Question C29.2: Who is right here?

When can we get away with computing $r(\mathbf{A})$ only?

Bob: Good point, Cindy! Recall from Lecture 27 that the set of all $m \times 1$ vectors $\vec{\mathbf{b}}$ such that the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ has a solution is the set $CS(\mathbf{A})$ of all linear combinations of the columns of $\mathbf{A}$. This is a linear subspace of $\mathbb{R}^m$ and always contains the zero vector $\vec{\mathbf{0}}$.

Theo: The dimension of $CS(\mathbf{A})$ is $r(\mathbf{A})$. When $r(\mathbf{A}) = m$, this must be the entire space $\mathbb{R}^m$; and when $r(\mathbf{A}) < m$, it will not be the entire space.

Alice: When $r(\mathbf{A}) < m$, for **some** $\vec{\mathbf{b}}$ the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ will be *inconsistent,* and for **some** $\vec{\mathbf{b}}$ the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ will be *consistent.*

Cindy: I see now how the rank will tell me whether the system has 0 solutions or at least 1 solution. But how can I tell whether it has **exactly 1** or **more than 1** solutions?

When does the system have exactly one solution?

Theo: If a system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ with $m \times n$ coefficient matrix $\mathbf{A}$ is consistent, you can tell this by comparing the rank $r(\mathbf{A})$ with the *number n of columns* of $\mathbf{A}$.

If $r(\mathbf{A}) = n$, then any consistent system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ will have exactly one solution.

Bob: Why is that?

Theo: Because when $r(\mathbf{A}) = n$, then the columns of $\mathbf{A}$ form a linearly independent set, which must be a basis of the column space $CS(\mathbf{A})$.

Then for every vector $\vec{\mathbf{b}}$ in the column space $CS(\mathbf{A})$ there exists **exactly one** vector vector of coefficients $\vec{\mathbf{x}} = [x_1, \dots, x_n]^T$ such that

$\vec{\mathbf{b}} = x_1\vec{\mathbf{a}}_{*1} + \cdots + x_n\vec{\mathbf{a}}_{*n}$, where $\vec{\mathbf{a}}_{*1}, \dots, \vec{\mathbf{a}}_{*n}$ are the columns of $\mathbf{A}$.

Question 29.3: How is this related to Cindy's question?

What if $r(\mathbf{A}) < n$?

Here $\vec{\mathbf{b}}$ would be in the column space $CS(\mathbf{A})$ if, and only if, the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is consistent. Moreover, $\vec{\mathbf{x}}$ would be a vector of coefficients for the linear combination if, and only if, $\vec{\mathbf{x}}$ is a solution of this system.

Denny: Would $r(\mathbf{A}) = n$ mean the same thing as saying that $\mathbf{A}$ has full rank?

Theo: Only if $\mathbf{A}$ is a square matrix. But $\mathbf{A}$ could have more rows than columns, and it would still be true that if $r(\mathbf{A}) = n$, then every consistent system has exactly one solution.

Cindy: But what if $\mathbf{A}$ has n columns and $r(\mathbf{A}) < n$?

Theo: Then the columns $\vec{\mathbf{a}}_{*1}, \dots, \vec{\mathbf{a}}_{*n}$ of $\mathbf{A}$ do not form a basis of $CS(\mathbf{A})$ and every system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is either inconsistent or has more than 1 solution.

So what if $r(\mathbf{A}) < n$?

Cindy: But how should I think about the solution set when it contains more than one vector? I always find this difficult!

Alice: Think about 2 solutions $\vec{\mathbf{x}}, \vec{\mathbf{y}}$ of the system, so that $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ and $\mathbf{A}\vec{\mathbf{y}} = \vec{\mathbf{b}}$.

Question C29.4: What do you get when you compute $\mathbf{A}(\vec{\mathbf{x}} - \vec{\mathbf{y}})$?

Cindy: I get $\mathbf{A}(\vec{\mathbf{x}} - \vec{\mathbf{y}}) = \mathbf{A}\vec{\mathbf{x}} - \mathbf{A}\vec{\mathbf{y}} = \vec{\mathbf{b}} - \vec{\mathbf{b}} = \vec{\mathbf{0}}$.

Alice: Right!

So $\vec{\mathbf{x}} - \vec{\mathbf{y}}$ is a solution of the homogeneous system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{0}}$.

Denny: Which means $\vec{\mathbf{x}} - \vec{\mathbf{y}}$ is in the null space $N(\mathbf{A})$ of $\mathbf{A}$. Cool!

Bob: I think it also should work the other way around: If $\vec{\mathbf{x}}$ is a solution of the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ and $\vec{\mathbf{z}}$ is in the null space $N(\mathbf{A})$ of $\mathbf{A}$, then $\vec{\mathbf{y}} = \vec{\mathbf{x}} + \vec{\mathbf{z}}$ is also a solution of this system.

Question C29.5: Does it work in the way that Bob suggested?

The null space and the solutions of $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$

Alice: Yes. I think Cindy can show us the proof.

Cindy: Suppose $\vec{\mathbf{x}}$ is a solution of the system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ and $\vec{\mathbf{z}}$ is in the null space $N(\mathbf{A})$ of $\mathbf{A}$. Then $\mathbf{A}(\vec{\mathbf{x}} + \vec{\mathbf{z}}) = \mathbf{A}\vec{\mathbf{x}} + \mathbf{A}\vec{\mathbf{z}} = \vec{\mathbf{b}} + \vec{\mathbf{0}} = \vec{\mathbf{b}}$. □

Alice: Thank you Cindy! You have proved the following result:

Theorem

Suppose $\mathbf{A}$ *is the coefficient matrix of a linear system* $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$.

Let $\vec{\mathbf{x}}$ *be a solution of this system and let* $\vec{\mathbf{y}}$ *be another vector. Then* $\vec{\mathbf{y}}$ *is also a solution of the same system if, and only if,* $\vec{\mathbf{x}} - \vec{\mathbf{y}}$ *is in* $N(\mathbf{A})$.

In other words, when $\vec{\mathbf{x}}$ *is one solution of the system* $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$, *then all other solutions must be of the form* $\vec{\mathbf{x}} + \vec{\mathbf{z}}$, *where* $\vec{\mathbf{z}}$ *is in* $N(\mathbf{A})$.

Cindy: Wow! But how does this help me with representing solution sets if there is more than one solution?

Bob: Congratulations, Cindy! We will learn about this in Lecture 29.

Take-home message

Consider a linear system $\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{b}}$ with coefficient matrix $\mathbf{A}$ of order $m \times n$.

- The system is consistent if, and only if, $r(\mathbf{A}) = r([\mathbf{A}, \vec{\mathbf{b}}])$.
- When $r(\mathbf{A}) = m$, the system is always consistent; when $r(\mathbf{A}) < m$, the system is consistent for some, but not for all choices of $\vec{\mathbf{b}}$.
- When $r(\mathbf{A}) = n$, the system is either inconsistent or has exactly one solution.
- When $r(\mathbf{A}) < n$ and the system is consistent, then
  - The system is underdetermined.
  - When $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ are two solutions, then $\vec{\mathbf{x}} - \vec{\mathbf{y}}$ is in the null space $N(\mathbf{A})$.
  - When $\vec{\mathbf{x}}$ is one solution of the system, then all other solutions are vectors of the form $\vec{\mathbf{x}} + \vec{\mathbf{z}}$, where $\vec{\mathbf{z}}$ is in $N(\mathbf{A})$.