MATH 423

Linear Algebra II

Lecture 16:
Rank of a matrix.
Systems of linear equations.
Reduced row echelon form.

Rank of a matrix

Definition. The **rank** of an $m \times n$ matrix $A \in \mathcal{M}_{m,n}(\mathbb{F})$ is the rank of the linear transformation $L_A : \mathbb{F}^n \to \mathbb{F}^m$ given by $L_A(\mathbf{x}) = A\mathbf{x}$.

$$\mathbf{y} = A\mathbf{x} \quad \Longleftrightarrow \quad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$\Longleftrightarrow \quad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

Theorem The rank of the matrix A is the dimension of its *column space*, i.e., a subspace of $\mathbb{F}^m$ spanned by its columns.

Let V_1, V_2, and V_3 be finite-dimensional vector spaces. Suppose that $L : V_1 \to V_2$ and $T : V_2 \to V_3$ are linear transformations.

Theorem $\operatorname{rank}(T \circ L) \leq \min\bigl(\operatorname{rank}(T), \operatorname{rank}(L)\bigr)$.

Proof: Since $(T \circ L)(\mathbf{x}) = T(L(\mathbf{x}))$ for any $\mathbf{x} \in V_1$, it follows that $\mathcal{R}(T \circ L) \subset \mathcal{R}(T)$. Then $\dim \mathcal{R}(T \circ L) \leq \dim \mathcal{R}(T)$, i.e., $\operatorname{rank}(T \circ L) \leq \operatorname{rank}(T)$.

By the Dimension Theorem,

$$\dim \mathcal{R}(L) + \dim \mathcal{N}(L) = \dim \mathcal{R}(T \circ L) + \dim \mathcal{N}(T \circ L) = \dim V_1.$$

Hence the inequality $\operatorname{rank}(T \circ L) \leq \operatorname{rank}(L)$ is equivalent to the inequality $\dim \mathcal{N}(T \circ L) \geq \dim \mathcal{N}(L)$.

Let $\mathbf{0}_i$ denote the zero vector in the space V_i, $1 \leq i \leq 3$. If $L(\mathbf{x}) = \mathbf{0}_2$ for some vector $\mathbf{x} \in V_1$, then

$$(T \circ L)(\mathbf{x}) = T(L(\mathbf{x})) = T(\mathbf{0}_2) = \mathbf{0}_3.$$

This means that $\mathcal{N}(L) \subset \mathcal{N}(T \circ L)$. Consequently, $\dim \mathcal{N}(L) \leq \dim \mathcal{N}(T \circ L)$.

Theorem 1 Let A and B be matrices such that the product AB is well defined. Then

$$\operatorname{rank}(AB) \leq \min\bigl(\operatorname{rank}(A), \operatorname{rank}(B)\bigr).$$

Proof: Since $(AB)\mathbf{x} = A(B\mathbf{x})$ for any column vector $\mathbf{x}$ of an appropriate dimension, we have $L_{AB} = L_A \circ L_B$. Therefore this theorem is a corollary of the theorem from the previous slide.

Theorem 2 Let $A \in \mathcal{M}_{m,n}(\mathbb{F})$. Then for any invertible matrices $B \in \mathcal{M}_{n,n}(\mathbb{F})$ and $C \in \mathcal{M}_{m,m}(\mathbb{F})$,

$$\operatorname{rank}(A) = \operatorname{rank}(AB) = \operatorname{rank}(CA) = \operatorname{rank}(CAB).$$

Proof: By Theorem 1, $\operatorname{rank}(AB) \leq \operatorname{rank}(A)$. On the other hand, $\operatorname{rank}(A) = \operatorname{rank}((AB)B^{-1}) \leq \operatorname{rank}(AB)$. Therefore $\operatorname{rank}(AB) = \operatorname{rank}(A)$. Similarly, $\operatorname{rank}(CA) \leq \operatorname{rank}(A)$ $= \operatorname{rank}(C^{-1}(CA)) \leq \operatorname{rank}(CA)$. Finally, $\operatorname{rank}(CAB) = \operatorname{rank}((CA)B) = \operatorname{rank}(CA) = \operatorname{rank}(A)$.

Corollary Elementary row and column operations do not change the rank of a matrix.

Proof: Elementary row (resp. column) operations can be simulated as left (resp. right) multiplication by the elementary matrices. Since the elementary matrices are invertible, such multiplication does not change the rank of a matrix.

Theorem $\operatorname{rank}(A^t) = \operatorname{rank}(A)$.

Proof: First we consider a special case when A is a block matrix of the form $\begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix}$, where I_r is the identity matrix of dimensions $r \times r$ and O_1, O_2, O_3 are zero matrices of appropriate dimensions. Namely, if A is $m \times n$, then O_1 is $r \times (n-r)$, O_2 is $(m-r) \times r$, and O_3 is $(m-r) \times (n-r)$. For example, in the case $r = 2$, $m = 3$, $n = 4$ we have

$$A = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right).$$

The first r columns of A are the first r vectors from the standard basis for $\mathbb{F}^m$, the other columns are zero vectors. It follows that $\operatorname{rank}(A) = r$. Furthermore, the first r columns of A^t (which are the first r rows of A) are the first r vectors from the standard basis for $\mathbb{F}^n$. The other columns of A^t are zero vectors. It follows that $\operatorname{rank}(A^t) = r$.

Theorem $\operatorname{rank}(A^t) = \operatorname{rank}(A)$.

Proof (continued): The general case will be reduced to the special case using the following lemma (which is a particular case of a theorem proved in Lecture 11).

Lemma There exists a basis α for $\mathbb{F}^n$ and a basis β for $\mathbb{F}^m$ such that the matrix $D = [L_A]_\alpha^\beta$ is of the form $\begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix}$.

Let γ denote the standard basis for $\mathbb{F}^n$ and δ denote the standard basis for $\mathbb{F}^m$. Let id_S denote the identity mapping on a set S. We have $L_A = \mathrm{id}_{\mathbb{F}^m} \circ L_A \circ \mathrm{id}_{\mathbb{F}^n}$, which implies $[L_A]_\alpha^\beta = [\mathrm{id}_{\mathbb{F}^m}]_\delta^\beta \, [L_A]_\gamma^\delta \, [\mathrm{id}_{\mathbb{F}^n}]_\alpha^\gamma$. Notice that $[L_A]_\gamma^\delta = A$, $[\mathrm{id}_{\mathbb{F}^n}]_\alpha^\gamma$ is the transition matrix from α to γ, and $[\mathrm{id}_{\mathbb{F}^m}]_\delta^\beta$ is the transition matrix from δ to β. Therefore $D = CAB$, where B and C are invertible matrices. By the previous theorem, $\operatorname{rank}(D) = \operatorname{rank}(A)$. Further, $D^t = (CAB)^t = B^t A^t C^t$. The transposes B^t and C^t are also invertible, hence $\operatorname{rank}(D^t) = \operatorname{rank}(A^t)$. Finally, $\operatorname{rank}(D^t) = \operatorname{rank}(D)$.

System of linear equations

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

Here $x_1, x_2, \ldots, x_n$ are variables and a_{ij}, b_j are constants.

A *solution* of the system is a common solution of all equations in the system. It is regarded an n-dimensional coordinate vector.

A system of linear equations can have **one** solution, **infinitely many** solutions, or **no** solution at all.

System of linear equations:

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

Coefficient matrix and column vector of the right-hand sides:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \qquad \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

System of linear equations:

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

Augmented matrix:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right)$$

Matrix representation of the system

$$\left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right. \iff A\mathbf{x} = \mathbf{b},$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Theorem The system is consistent (i.e., has a solution) if and only if $\mathrm{rank}(A) = \mathrm{rank}(A \,|\, \mathbf{b})$.

Row echelon form

Definition. **Leading entry** of a matrix is the first nonzero entry in a row.

A matrix is said to be in the **row echelon form** if the leading entries shift to the right as we go from the first row to the last one.

Example.

$$\left(\begin{array}{cccccccccc|c}
1 & -1 & 3 & 0 & 2 & 1 & 4 & 0 & -3 & 7 & 2 \\
0 & 1 & 1 & -2 & 0 & 0 & 6 & 1 & 3 & 4 & 0 \\
0 & 0 & 0 & 0 & 1 & 2 & 3 & 1 & -4 & 2 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 9 & -1 & 2 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & -3 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right)$$

Row echelon form

General matrix in row echelon form:

$$\left(\begin{array}{cccccccccc|c}
\boxed{} & * & * & * & * & * & * & * & * & * & * \\
 & \boxed{} & * & * & * & * & * & * & * & * & * \\
 & & & & \boxed{} & * & * & * & * & * & * \\
 & & & & & & \boxed{} & * & * & * & * \\
 & & & & & & & \boxed{} & * & * & * \\
 & & & & & & & & & &
\end{array}\right)$$

- leading entries are boxed;
- all the entries below the staircase line are zero;
- each step of the staircase has height 1;
- each circle marks a column without a leading entry (for the augmented matrix of a system of linear equations, they will correspond to free variables).

Reduced row echelon form

A matrix is said to be in the **reduced row echelon form** if
(i) it is in the row echelon form (i.e., leading entries shift to the right as we go from the first row to the last one);
(ii) each leading entry is equal to 1;
(iii) each leading entry is the only nonzero entry in its column.

$$\left(\begin{array}{cccccccccc|c}
\boxed{1} & & * & * & & * & & & * & * & * \\
 & \boxed{1} & * & * & & * & & & * & * & * \\
 & & & & \boxed{1} & * & & & * & * & * \\
 & & & & & & \boxed{1} & & * & * & * \\
 & & & & & & & \boxed{1} & * & * & * \\
 & & & & & & & & & & \\
 & & & & & & & & & &
\end{array}\right)$$

- all entries below the staircase line are zero;
- each boxed entry is 1, the other entries in its column are 0.