Lecture 19

Rank

Definition 22.1: For any matrix A, put

$$\begin{aligned}\text{rank}(A) &= \dim(\text{img}(A)) \\ &= \dim(\text{ span of the columns of } A) \\ &= \dim(\text{ span of the rows of } A^T). \\ \text{nullity}(A) &= \dim(\ker(A)) \\ &= \dim(\text{ annihilator of the rows of } A)\end{aligned}$$

- $\text{rank}(A) = 0$ if and only if $A = 0$.
- $\text{rank}(A) = 1$ if and only if $A \neq 0$ and all nonzero columns are multiples of each other.
- An $n \times n$ matrix has rank n if and only if $\text{img}(A) = \mathbb{R}^n$, if and only if A is invertible.

Column operations

Definition 22.2: A matrix A is in *reduced column echelon form* (RCEF) if A^T is in RREF, or equivalently:

RCEF0: Any column of zeros come at the right hand end of the matrix, after all the nonzero columns.

RCEF1: In any nonzero column, the first nonzero entry is equal to one. These entries are called *copivots*.

RCEF2: In any nonzero column, the copivot is further down than the copivots in all previous rows.

RCEF3: If a row contains a copivot, then all other entries in that row are zero.

Definition 22.3: Let A be a matrix. The following operations on A are called *elementary column operations*:

ECO1: Exchange two columns.

ECO2: Multiply a column by a nonzero constant.

ECO3: Add a multiple of one column to another column.

Rank of an RCEF matrix

Proposition 22.4: If a matrix A is in RCEF, then the rank of A is just the number of nonzero columns.

Proof.
Let the nonzero columns be $u_1, \dots, u_r$, and put $U = \text{span}(u_1, \dots, u_r)$.
This is the same as the span of *all* the columns, because columns of zeros do not contribute anything to the span.
We claim that the vectors u_i are linearly independent.
To see this, note that each u_i contains a copivot, say in the q_i'th row. As the matrix is in RCEF we have $q_1 < \cdots < q_r$, and the q_i'th row is all zero apart from the copivot in u_i. In other words, for $j \neq i$ the q_i'th entry in u_j is zero.
Now suppose we have a linear relation $\lambda_1 u_1 + \cdots + \lambda_r u_r = 0$.
By looking at the q_i'th entry, we see that λ_i is zero.
This holds for all i, so we have the trivial linear relation.
This proves that the list $u_1, \dots, u_r$ is linearly independent, so it forms a basis for U, so $\dim(U) = r$. We thus have $\text{rank}(A) = r$ as claimed. □

Basic facts about column operations

Proposition 22.5: Any matrix A can be converted to RCEF by a sequence of elementary column operations.

Proof: Analogous to Method 6.3 for row operations. □

Proposition 22.6: Suppose that A can be converted to B by a sequence of elementary column operations. Then $B = AV$ for some invertible matrix V.

Proof.
A^T can be converted to B^T by a sequence of row operations corresponding to the column operations that were used to convert A to B.
Thus, Corollary 11.10 tells us that $B^T = UA^T$ for some invertible matrix U.
We thus have $B = B^{TT} = (UA^T)^T = A^{TT}U^T = AU^T$.
Here U^T is also invertible, so we can take $V = U^T$. □

Proposition 22.7: Suppose that A can be converted to B by a sequence of elementary column operations. Then the span of the columns of A is the same as the span of the columns of B (and so $\text{rank}(A) = \text{rank}(B)$).

Proof: Analogous to Corollary 9.16 for row operations. □

Invariance under row operations

Proposition 22.8: Suppose that A can be converted to B by a sequence of elementary *row* operations. Then $\text{rank}(A) = \text{rank}(B)$.

Proof: Let the columns of A be $v_1, \dots, v_n$ and put $V = \text{span}(v_1, \dots, v_n)$ so $\text{rank}(A) = \dim(V)$. There is an invertible matrix P such that

$$B = PA = P\left[\begin{array}{c|c|c} v_1 & \cdots & v_n \end{array}\right] = \left[\begin{array}{c|c|c} Pv_1 & \cdots & Pv_n \end{array}\right],$$

so the vectors Pv_i are the columns of B. Thus, if we put $W = \text{span}(Pv_1, \dots, Pv_n)$, then $\text{rank}(B) = \dim(W)$.

Claim: if $x \in V$ then $Px \in W$. Indeed, if $x \in V$ then $x = \sum_{i=1}^n \lambda_i v_i$ for some sequence of coefficients $\lambda_1, \dots, \lambda_n$. This means that $Px = \sum_{i=1}^n \lambda_i Pv_i$, which is a linear combination of $Pv_1, \dots, Pv_n$, so $Px \in W$.

Claim: if $y \in W$ then $P^{-1}y \in V$. Indeed, if $y \in W$ then $y = \sum_{i=1}^n \lambda_i Pv_i$ for some sequence of coefficients $\lambda_1, \dots, \lambda_n$. This means that $P^{-1}y = \sum_{i=1}^n \lambda_i P^{-1}Pv_i = \sum_{i=1}^n \lambda_i v_i$, which is a linear combination of $v_1, \dots, v_n$, so $P^{-1}y \in V$.

Invariance under row operations

$V = \text{span}(v_1, \dots, v_n)$; $W = \text{span}(Pv_1, \dots, Pv_n)$;
if $x \in V$ then $Px \in W$; if $y \in W$ then $P^{-1}y \in V$.

Now choose a basis $a_1, \dots, a_r$ for V (so $\text{rank}(A) = \dim(V) = r$).
Claim: the vectors $Pa_1, \dots, Pa_r$ form a basis for W.
We just showed that $Px \in W$ whenever $x \in V$, so at least $Pa_i \in W$.

Consider an arbitrary element $y \in W$. We then have $P^{-1}y \in V$, but the vectors a_i form a basis for V, so we have $P^{-1}y = \sum_{i=1}^r \mu_i a_i$ for some sequence of coefficients μ_i. This means that $y = PP^{-1}y = \sum_i \mu_i Pa_i$, which expresses y as a linear combination of the vectors Pa_i. It follows that the list $Pa_1, \dots, Pa_r$ spans W.

We need to check that it is also linearly independent.
Suppose we have a linear relation $\sum_i \lambda_i Pa_i = 0$. After multiplying by P^{-1}, we get a linear relation $\sum_i \lambda_i a_i = 0$. The list $a_1, \dots, a_r$ is assumed to be a basis for V, so this must be the trivial relation, so $\lambda_1 = \dots = \lambda_r = 0$, or in other words the original relation $\sum_i \lambda_i Pa_i = 0$ was the trivial one.

We have now shown that $Pa_1, \dots, Pa_r$ is a basis for W, so $\dim(W) = r$. In conclusion, we have $\text{rank}(A) = r = \text{rank}(B)$ as required. □

Normal form

Definition 22.9: An $n \times m$ matrix A is in *normal form* if it has the form

$$A = \left[\begin{array}{c|c} I_r & 0_{r\times(m-r)} \\ \hline 0_{(n-r)\times r} & 0_{(n-r)\times(m-r)} \end{array}\right]$$

for some r. ($r = 0$ is allowed, in which case A is just the zero matrix.)

If A is in normal form as above, then $\text{rank}(A) = r =$ the number of ones in A.

Example 22.10: There are precisely four different 3×5 matrices that are in normal form, one of each rank from 0 to 3 inclusive.

$$A_0 = \begin{bmatrix} 0&0&0&0&0 \\ 0&0&0&0&0 \\ 0&0&0&0&0 \end{bmatrix} \qquad A_1 = \left[\begin{array}{c|cccc} 1&0&0&0&0 \\ \hline 0&0&0&0&0 \\ 0&0&0&0&0 \end{array}\right]$$

$$A_2 = \left[\begin{array}{cc|ccc} 1&0&0&0&0 \\ 0&1&0&0&0 \\ \hline 0&0&0&0&0 \end{array}\right] \qquad A_3 = \left[\begin{array}{ccc|cc} 1&0&0&0&0 \\ 0&1&0&0&0 \\ 0&0&1&0&0 \end{array}\right]$$

Reduction to normal form

Proposition 22.11: Any $n \times m$ matrix A can be converted to a matrix C in normal form by a sequence of row and column operations. Moreover:

(a) There is an invertible $n \times n$ matrix U and an invertible $m \times m$ matrix V such that $C = UAV$.

(b) $\text{rank}(A) = \text{rank}(C) =$ the number of ones in C.

Proof: Perform row operations to get a matrix B in RREF.
By Corollary 11.10 there is an invertible matrix U such that $B = UA$.
(This has to be an $n \times n$ matrix for the product UA to make sense.)
Now subtract multiples of pivot columns from columns further to the right.
As each pivot column contains nothing but the pivot, the only effect of these column operations is to set everything to the right of a pivot equal to zero.
However, every nonzero entry in B is either a pivot or to the right of a pivot, so after these ops we just have the pivots from B and everything else is zero.
Now just move all columns of zeros to the right hand end, which leaves a matrix C in normal form. As C was obtained from B by a sequence of elementary column operations, we have $C = BV$ for some invertible $m \times m$ matrix V. As $B = UA$, it follows that $C = UAV$. Propositions 22.7 and 22.8 tell us that neither row nor column operations affect the rank, so $\text{rank}(A) = \text{rank}(C)$, and because C is in normal form, $\text{rank}(C)$ is just the number of ones in C.

Example of reduction to normal form

Consider the matrix

$$A = \begin{bmatrix} 1 & 3 & 0 & 1 \\ 2 & 6 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 1 & 3 & 2 & 9 \end{bmatrix}.$$

This can be row-reduced as follows:

$$\begin{bmatrix} 1 & 3 & 0 & 1 \\ 2 & 6 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 1 & 3 & 2 & 9 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 2 & 8 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 0 & 1 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We now perform column operations:

$$\begin{bmatrix} 1 & 3 & 0 & 1 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

(Subtract column 1 from column 4, and 3 times column 1 from column 2; subtract 4 times column 3 from column 4; exchange columns 2 and 3.)
We are left with a matrix of rank 2 in normal form, so $\text{rank}(A) = 2$.

Example of reduction to normal form

Consider the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix}$$

This can be reduced to normal form as follows: $A \to$

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & -2 & -4 \\ 0 & -3 & -6 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -2 & -4 \\ 0 & -3 & -6 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

- Subtract multiples of row 1 from the other rows
- Multiply row 2 by -1
- Subtract multiples of row 2 from the other rows
- Add column 1 to column 3
- Subtract 2 times column 2 from column 3.

The final matrix has rank 2, so we must also have $\text{rank}(A) = 2$.

$\text{rank}(A) = \text{rank}(A^T)$

Proposition 22.14: For any matrix A we have $\text{rank}(A) = \text{rank}(A^T)$.

Proof: We can convert A by row and column ops to a matrix C in normal form, and $\text{rank}(A)$ is the number of ones in C. If we transpose everything then the row ops become column ops and *vice-versa*, so A^T can be converted to C^T by row and column ops, and C^T is also in normal form, so $\text{rank}(A^T) =$ the number of ones in $C^T =$ the number of ones in $C = \text{rank}(A)$. □

Alternative terminology:

column rank of $A = \dim(\text{span}(\text{ columns of } A)) = \text{rank}(A)$

row rank of $A = \dim(\text{span}(\text{ rows of } A)) = \dim(\text{span}(\text{ cols of } A^T)) = \text{rank}(A^T)$

With this terminology, the proposition says row rank=column rank.

Corollary 22.16: If A is an $n \times m$ matrix. Then $\text{rank}(A) \leq \min(n, m)$.

Proof: Let V be the span of the columns of A, and let W be the span of the columns of A^T. Now V is a subspace of $\mathbb{R}^n$, so $\dim(V) \leq n$, but W is a subspace of $\mathbb{R}^m$, so $\dim(W) \leq m$. On the other hand, Proposition 22.14 tells us that $\dim(V) = \dim(W) = \text{rank}(A)$, so we have $\text{rank}(A) \leq n$ and also $\text{rank}(A) \leq m$, so $\text{rank}(A) \leq \min(n, m)$.