1. Rank of a matrix

Definition 1.1. Let $A = (a_{ij})_{1\le i\le m, 1\le j\le n}$ be an $m \times n$ matrix. An $m' \times n'$ submatrix B is obtained by choosing m' rows and n' columns and removing all entries outside of these rows and columns. Formally speaking, one fixes numbers $1 \le i_1 < i_2 < \cdots < i_{m'} \le m$ and $1 \le j_1 < j_2 < \cdots < j_{n'} \le n$ and defines $B = (b_{pq})_{1\le p\le m', 1\le q\le n'}$ by $b_{pq} = a_{i_p j_q}$. We say that B is the submatrix whose rows are $i_1, i_2, \dots, i_{m'}$ and columns are $j_1, j_2, \dots, j_{n'}$.

Definition 1.2. Given an $m \times n$ matrix A one defines the following numbers:

(i) The *rank* of A is the maximal number $r = \mathrm{rk}(A)$ such that A possesses an invertible $r \times r$ submatrix.

(ii) The *row rank* of A is the maximal number $r = \mathrm{rk}_{\mathrm{row}}(A)$ such that A possesses r linearly independent rows (in the space of rows F^n).

(iii) The *column rank* $\mathrm{rk}_{\mathrm{col}}(A)$ is defined similarly to $\mathrm{rk}_{\mathrm{row}}(A)$ but using columns instead of rows.

Lemma 1.3. *Let A be a matrix, then*

(i) $\mathrm{rk}(A) = \mathrm{rk}(A^t)$*,*

(ii) $\mathrm{rk}_{\mathrm{row}}(A) = \mathrm{rk}_{\mathrm{col}}(A^t)$*.*

Proof. Observe that the rows of A are the columns of A^t and vice versa, so we obtain (ii). In addition, this observation implies that if B is a submatrix of A given by rows $i_1, \dots, i_{m'}$ and columns $j_1, \dots, j_{n'}$ then B^t is the submatrix of A^t given by rows $j_1, \dots, j_{n'}$ and columns $i_1, \dots, i_{m'}$. Since B is invertible if and only if B^t is invertible, we obtain that $\mathrm{rk}(A) = \mathrm{rk}(A^t)$. □

Lemma 1.4. *If $v_1 = (x_{11}, \dots, x_{1n}), \dots, v_r = (x_{r1}, \dots, x_{rn})$ are r linearly independent vectors in F^n and $n > r$, then there exists $1 \le j \le n$ such that the vectors $w_1 = (x_{11}, \dots, \widehat{x_{1j}}, \dots, x_{1n}), \dots, w_r = (x_{r1}, \dots, \widehat{x_{rj}}, \dots, x_{rn})$ are linearly independent in F^{n-1}.*

Proof. Since the vectors are linearly independent, $V = \mathrm{Span}(v_1, \dots, v_r)$ is of dimension r. In particular, $V \ne F^n$, and there exists a standard basis vector ε_j not contained in V. Consider the linear map (in fact, a projection) $p\colon F^n \to F^{n-1}$ given by $p((x_1, \dots, x_n)) = (x_1, \dots, \widehat{x_j}, \dots, x_n)$. Its kernel consists of all vectors with $x_k = 0$ for any $k \ne j$, so $\mathrm{Ker}(p) = \mathrm{Span}(\varepsilon_j)$. Let $q\colon V \to F^{n-1}$ be the restriction of p onto V, i.e., $q\colon V \to F^{n-1}$ is the linear map given by $q(v) = p(v)$ for $v \in V$. Since $\varepsilon_j \notin V$, we have that $\mathrm{Ker}(q) = \mathrm{Ker}(p) \cap V = \mathrm{Span}(\varepsilon_j) \cap V = 0$. Thus, q is an embedding and $\dim(\mathrm{Im}(q)) = r$. But $\mathrm{Im}(q)$ is generated by r vectors $q(v_i) = (x_{i1}, \dots, \widehat{x_{ij}}, \dots, x_{in})$ with $1 \le i \le r$, so these vectors are linearly independent and we are done. □

Theorem 1.5. *For any matrix A all three ranks coincide:* $\mathrm{rk}(A) = \mathrm{rk}_{\mathrm{row}}(A) = \mathrm{rk}_{\mathrm{col}}(A)$*.*

Proof. We start with the equation $\mathrm{rk}(A) = \mathrm{rk}_{\mathrm{row}}(A)$. Assume that B is an $r \times r$ matrix given by rows $i_1, \dots, i_r$ and columns $j_1, \dots, j_r$. If B is invertible then its rows are linearly independent (by our theory of invertible matrices). It follows that the rows $i_1, i_2, \dots, i_r$ of A are also linearly independent, hence $\mathrm{rk}_{\mathrm{row}}(A) \ge r = \mathrm{rk}(A)$.

Conversely, assume that $r = \mathrm{rk}_{\mathrm{row}}(A)$ and choose r linearly independent rows of A, say $v_1 = (a_{i_1 1}, \dots, a_{i_1 n}), \dots, v_r = (a_{i_r 1}, \dots, a_{i_r n})$. The rows live in the n-dimensional row space, so $r \le n$. If $r < n$ then by the above lemma there exists

$1 \leq j \leq n$ such that the rows remain linearly independent after removing the j-th column. So, we can remove columns one by one obtaining a chain of submatrixes $r \times n, r \times (n-1), \ldots, r \times r$ such that each submatrix has linearly independent rows. The last submatrix is a square matrix, so linear independence of its rows implies invertibility (by the theory of invertible matrices). We found an $r \times r$ invertible submatrix, so $\mathrm{rk}(A) \geq r = \mathrm{rk}_{\mathrm{row}}(A)$.

The two inequalities imply that $\mathrm{rk}(A) = \mathrm{rk}_{\mathrm{row}}(A)$ for any matrix A. In particular, $\mathrm{rk}(A^t) = \mathrm{rk}_{\mathrm{row}}(A^t)$ and by the first lemma we obtain that $\mathrm{rk}_{\mathrm{col}}(A) = \mathrm{rk}_{\mathrm{row}}(A^t) = \mathrm{rk}(A^t) = \mathrm{rk}(A)$, completing the proof. □