Appendix A: The Rank of a Matrix

There must be a thousand theorems about matrices, and no doubt there are good texts that cover them all in a logical order. Here we examine just a tiny piece of this large pie. We wish to establish a few facts about the rank of a matrix. But some framework is needed in order to do this. We will carry out our task with a few "definitions" and "theorems". We assume the reader is aware of basic matrix facts such as the fact that adding a multiple of one column of a square matrix to another column does not change the determinant of the matrix.

A.1 Linear dependence and linear independence

This topic is slightly more subtle that one at first expects.

Definitions: Consider a set of n vectors $\mathbf{v_i}$ each of which has m components. This **set** of vectors $\{\mathbf{v_i}\}$ is said to be **linearly independent** if none of the vectors can be written as a linear combination of the other vectors. For an arbitrary given set $\{\mathbf{v_i}\}$ it is always possible to find some number N of these vectors which are linearly independent, and then the rest of the vectors can be written as linear combinations of those N independent vectors. If N = n, then the entire set $\{\mathbf{v_i}\}$ is linearly independent. If N < n, then we can write N = n-s for some s>0. In this case, we have N linearly independent vectors in our set, and in addition we have s **linearly dependent vectors**, since each of those s vectors can be written as a linear combination of the independent vectors. In this case one says that the entire **set** $\{\mathbf{v_i}\}$ is **linearly dependent** since it is not linearly independent. Since the vectors have m components, we know that there cannot be more than m linearly independent vectors in the set. We would have m independent vectors for example if the set $\{\mathbf{v_i}\}$ contained the m unit vectors. So $0 \le N \le m$. If all the vectors are zero vectors, N = 0. This number N of independent vectors has no commonly used name (as far as we know), but in the matrix context it will acquire a name.

Example 1: Consider $\{\mathbf{v_i}\}$ to be the following set of n = 3 unit vectors each with m = 3 components:

$$\begin{pmatrix}1\\0\\0\end{pmatrix} \quad \begin{pmatrix}0\\1\\0\end{pmatrix} \quad \begin{pmatrix}0\\0\\1\end{pmatrix} \qquad N = 3 \quad s = 0 \tag{A.1.1}$$

Since none of these vectors can be written as a linear combination of the others, this is a linearly independent set of vectors, and N = n = 3.

Example 2: Consider $\{\mathbf{v_i}\}$ to be the following set of n = 3 vectors each with m = 3 components:

$$\begin{pmatrix}1\\0\\0\end{pmatrix} \quad \begin{pmatrix}0\\1\\0\end{pmatrix} \quad \begin{pmatrix}1\\2\\0\end{pmatrix} \qquad N = 2 \quad s = 1 \tag{A.1.2}$$

We can take the first two vectors to be a set of linearly independent vectors and then the third vector can be written as a linear combination of the first two. Thus, in this case the full set of n = 3 vectors is linearly dependent. The number of independent vectors is N = 2, and the number of dependent vectors is s = 1.

Given the full set of n vectors, the subset of independent vectors is not unique. In Example 2 above one could have taken the first and third vectors as the independent pair.

Example 3: Consider $\{\mathbf{v}_i\}$ to be the following set of n = 3 vectors each with m = 3 components:

$$\begin{pmatrix}1\\0\\0\end{pmatrix} \quad \begin{pmatrix}0\\0\\0\end{pmatrix} \quad \begin{pmatrix}0\\0\\0\end{pmatrix} \qquad N = 1 \quad s = 2 \tag{A.1.3}$$

In this case, there is only N = 1 independent vector. The other two vectors can be written as "linear combinations" of the independent vectors in this sense: $\mathbf{v}_2 = 0 * \mathbf{v}_1$ and $\mathbf{v}_3 = 0 * \mathbf{v}_1$. Thus, the two all-zero vectors are classified as dependent vectors.

Suppose in a set of n vectors the vectors $\mathbf{v}_1$, $\mathbf{v}_2$....$\mathbf{v}_N$ are linearly independent, and the remaining vectors $\mathbf{v}_{N+1}$, $\mathbf{v}_{N+2}$..... $\mathbf{v}_{N+s}$ are linearly dependent. One can then write,

$$\begin{aligned}
\mathbf{v}_{N+1} &= \Sigma_{i=1}^{N} a^{(N+1)}{}_i \mathbf{v}_i \\
\mathbf{v}_{N+2} &= \Sigma_{i=1}^{N} a^{(N+2)}{}_i \mathbf{v}_i \\
&\cdots \\
\mathbf{v}_{N+s} &= \Sigma_{i=1}^{N} a^{(N+s)}{}_i \mathbf{v}_i
\end{aligned} \tag{A.1.4}$$

where in each equation $a^{(N+j)}{}_i$ is some set of constants describing the linear combination. If the dependent vector $\mathbf{v}_{N+2}$ were all zeros, its equation would just be $\mathbf{v}_{N+2} = 0$ and all constants $a^{(N+2)}{}_i$ vanish.

If the set of n vectors is linearly dependent, there exists at least one equation of the form shown in (A.1.4) above. If we pick any of these equations (there may be only one), it has the general form

$$\Sigma_{i=1}^{n} k_i \ \mathbf{v}_i = 0. \tag{A.1.5}$$

where the set of n coefficients $\{k_i|$ has at least one non-zero element. For example, suppose s = 1 and there is only one dependent vector with the first equation shown in (A.1.4). Then one can write

$$-\mathbf{v}_{N+1} + \Sigma_{i=1}^{N} a^{(N+1)}{}_i \mathbf{v}_i = 0 \tag{A.1.6}$$

and this has the form (A.1.5) where $k_i = a^{(N+1)}{}_i$ for i = 1,2..N and k_i= -1 for i = N+1. If it happened that this vector $\mathbf{v}_{N+1}$ were an all-zero vector, one would have $k_i = a^{(N+1)}{}_i = 0$ for i = 1,2..N and k_i= -1 for i = N+1. In this case the set $\{k_i\}$ has one non-zero element which is -1.

Notice that if we take $k_i \rightarrow 2\ k_i$ in (A.1.5) the equation is still true, so the overall scale of the set $\{k_i\}$ is undetermined. One could select any non-zero element of the set $\{k_i\}$ and set it to 1 and thus establish the scale of the remaining non-zero k_i.

Fact: If a set of n vectors is linearly dependent, there exists at least one equation of the form (A.1.5) where at least one element of the set $\{k_i\}$ is non-zero. (A.1.7)

A.2 A set of vectors interpreted as a matrix: column rank and row rank

Fact: One can regard an m x n matrix A (m rows, n columns) as a set of n column vectors $\mathbf{c}_i$ each of which has m components,

$$A = (\mathbf{c}_1, \mathbf{c}_2. \ldots \mathbf{c}_n) \ . \tag{A.2.1}$$

Some number N of these column vectors will be linearly independent. N certainly cannot exceed n. Since the vectors each have m components, N cannot exceed m. Thus, $N \leq n$ and $N \leq m$. One would express this by saying $N \leq \min(n,m)$.

Definition: When the set of n vectors is considered to be the set of columns of a matrix A, this number N which is the number of linearly independent column vectors has a name. N is called the **column rank** of matrix A. Thus we have

$$0 \leq \text{column rank} \leq \min(n,m). \tag{A.2.2}$$

Definition: One can instead represent the matrix A in terms of m vectors $\mathbf{r}_i$ which are the rows of A,

$$A = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \ldots \\ \mathbf{r}_m \end{pmatrix} \tag{A.2.3}$$

In this case, some number N' of these vectors will be linearly independent. Thus number N' is called the **row rank** of the matrix A. By the same argument presented above, we must have

$$0 \leq \text{row rank} \leq \min(n,m). \tag{A.2.4}$$

Comment: We will show below in (A.4.8) that for any mxn matrix A, the column rank and the row rank must be the same, and thus one can just talk about the "rank" of a matrix. But since we have not yet demonstrated this fact, we continue to use column rank and row rank as separate entities. The rank of a matrix is easily determined by reducing the matrix to "row echelon form" using "elementary row operations", see for example the wiki rank (linear algebra) page or any matrix text.

A.3 Linear transformations: domain, range, dimension, nullspace and nullity

In this subsection and those that follow we ignore technical details (such as the definition of a "subspace" and the meaning of "span") and work with a low level of mathematical rigor, just to get things laid out without extra baggage. The reader can find rigorous forms of the ideas below in texts on matrix theory or linear algebra.

Definition: A linear transformation $A : E^n \rightarrow E^m$ can be represented as $A\mathbf{x} = \mathbf{y}$ where $\mathbf{x}$ is a vector in E^n having n components, and where $\mathbf{y}$ is a vector in E^m having m components. The matrix A will have a number of columns equal to the number of components of $\mathbf{x}$, which is n. The matrix will have a number of rows which is equal to the number of components of $\mathbf{y}$, which is m. We shall allow $\mathbf{x}$ to be any vector in the space E^n, so the **domain** of the transformation is all of E^n . The vectors $\mathbf{y}$ lie in E^m, so we might refer to E^m as the "range space" for the transformation. However, as $\mathbf{x}$ ranges over all of E^n, $\mathbf{y}$ might range only over a subset of the range space, and that subset is known as the **range** of the transformation. In our case then we would say domain = E^n while range $\subseteq E^m$ (is contained in E^m). The domain and range each have a **dimension** which is the number of independent vectors it takes to "span" the domain or range. In our case, the dimension of the domain D = E^n is clearly n. The dimension of the range $R \subseteq E^m$ is less than or equal to m.

Theorem. In the linear transformation context $A : E^n \rightarrow E^m$, the dimension of the range of the transformation is equal to the column rank of the mxn matrix A.

$$\dim(R) = \text{column rank}(A) \qquad \text{(A.3.1)}$$

Recall from (A.2.2) that column rank ≤ m, and m is certainly the maximum dimension the range $R \subseteq E^m$ could have.

Proof: Write the matrix equation $A\mathbf{x} = \mathbf{y}$ in this manner using (A.2.1) that $A = (\mathbf{c}_1, \mathbf{c}_2. \ldots \mathbf{c}_n)$:

$$(\mathbf{c}_1, \mathbf{c}_2. \ldots \mathbf{c}_n) \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_m \end{pmatrix} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \ldots\ldots x_n\mathbf{c}_n \ . \qquad \text{(A.3.2)}$$

Suppose the column rank of matrix A is N, so there are at most N linearly independent $\mathbf{c}_i$ vectors. We can then write

$$(\mathbf{c}_1, \mathbf{c}_2. \ldots \mathbf{c}_n) \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_m \end{pmatrix} = (x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \ldots\ldots x_N\mathbf{c}_N) + (x_{N+1}\, \mathbf{c}_{N+1} \ldots\ldots x_n\mathbf{c}_n) \qquad \text{(A.3.3)}$$

Since each of the vectors in the second group $\mathbf{c}_{N+1}$ to $\mathbf{c}_n$ can be written as a linear combination of the vectors in the first group, as in (A.1.4), if we replace each of these vectors with its linear combination, we end up with something of this form

$$(\mathbf{c}_1, \mathbf{c}_2. \ldots \mathbf{c}_n) \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_m \end{pmatrix} = (k_1\mathbf{c}_1 + k_2\mathbf{c}_2 + \ldots\ldots k_N\mathbf{c}_N) \ . \qquad \text{(A.3.4)}$$

This says that the range is spanned by N vectors $\mathbf{c}_1$ through $\mathbf{c}_N$ and thus the dimension of the range is N, which is the column rank of A. **QED**.

Definition. In our linear transformation A : $E^n \rightarrow E^m$ of the form A**x** = **y** , it may happen that for a subset $\mathcal{N}$ of the domain D = E^n, we find that if **x** is in $\mathcal{N}$, then A**x** = 0. This subset $\mathcal{N}$ (a subspace of E^n) is called the **nullspace** of the transformation A. It is that space within the domain that maps into the null vector 0 in the range. Like the domain and range, the nullspace has some dimension which is the number of independent vectors it takes to span the nullspace. The vector 0 always maps into 0, and if this is the only vector that does so, then the nullspace consists of just the vector 0 and the nullspace has dimension 0. This is sometimes called "the trivial nullspace" and is of no real interest. The dimension of the nullspace has a special name: it is called the **nullity** of A. So then

$$\text{nullity(A)} = \dim(\mathcal{N}) = \text{the dimension of the nullspace of A} \qquad \text{(A.3.5)}$$

Theorem. Consider a matrix A with m rows and n columns. Then

$$\text{nullity(A)} + \text{column rank(A)} = n. \quad \text{OR} \quad \dim(\mathcal{N}) + \dim(R) = n \qquad \text{(A.3.6)}$$

Proof: We can partition the domain in this way : D = $\mathcal{N} \cup \mathcal{N}^{\perp}$ where $\mathcal{N}^{\perp}$ (the perp space) is just the rest of the domain when the nullspace is removed. We must have dim(D) = dim($\mathcal{N}$) + dim($\mathcal{N}^{\perp}$), which we can write as n = nullity(A) + dim($\mathcal{N}^{\perp}$) . Since every **x** in $\mathcal{N}^{\perp}$ maps into the range of the transformation, it must be that dim($\mathcal{N}^{\perp}$) = dim(R). But from Theorem (A.3.1), we know that dim(R) = column rank(A). Therefore we have shown that n = nullity(A) + column rank (A). QED.

If one studies a large set of random mxn matrices A, one always finds that the sum of the dimension of the nullspace plus the dimension of the range equals the number n of columns of A. Some matrices have a large nullspace and a small range, other matrices have a small nullspace and a large range.

A.4 Some Matrix Theorems

Theorem 1A. For an n x n square matrix A, if det(A) ≠ 0, then the columns are linearly indenep (A.4.1)

Contrapositive: column rank < n ⇒ det(A) = 0.

Proof of the contrapositive: If the column rank is less than the full n, the columns are linearly dependent and one can write $\Sigma_i k_i \mathbf{c}_i = 0$ where at least one k_j does not vanish. Then $\mathbf{c}_j = -\Sigma_{i \neq j} (k_i/k_j)\mathbf{c}_i$. We know that adding a multiple of column I to column J does not change det(A). So det(A) does not change if we add $-\Sigma_{i \neq j} (k_i/k_j)\mathbf{c}_i$ to $\mathbf{c}_j$ which makes the new column vector $\mathbf{c}_j = 0$. A matrix with a zero column has 0 determinant, so det(A) = 0.

We now consider the converse of the above theorem. We know that N ≡ column rank(A) ≤ n since this N is the number of independent columns of A. If it happens that column rank(A) = n, then the square matrix has **full rank**. It is a very familiar fact that in this case det(A) ≠ 0, but why is this so?

Theorem. For an n x n square matrix A, if the column rank is n, then det(A) ≠ 0. (A.4.2)

Proof #1. If the n columns of A are linearly independent, then each column must contain at least one non-zero number. One cannot have a column of all zeros. We know we can swap rows around without changing the value of |det(A)|. So find a row which has a non-zero number in the left column and swap this with the first row so there is a non-zero number in the upper left corner. Then add multiples of this new first row to all the other rows so as to clear out (to zeros) the rest of the first column. These operations also do not affect det(A).

After

After doing the above, find a non-zero number in the second column (it must exist WHY?) and swap rows so it is in the A_{22} matrix position.

Repeat this process to get a matrix whose lower left triangle is all zeros and whose diagonal has all non-zero values. This matrix then has upper triangular form. The determinant of such a matrix is the product of the diagonal values and thus we have shown that |det(A)| ≠ 0. This process is usually called Gauss elimination. One might wonder why *all* the diagonal elements are non-zero for some general case. This is assured at each step of the elimination. We leave it to the reader to make this proof more rigorous.

Example: (The row swap negates the determinant:)

$$A = \begin{pmatrix} 0 & 1 & 3 \\ 6 & 0 & 4 \\ 2 & 5 & 7 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 5 & 7 \\ 6 & 0 & 4 \\ 0 & 1 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 5 & 7 \\ 0 & -15 & -17 \\ 0 & 1 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 5 & 7 \\ 0 & -15 & -17 \\ 0 & 0 & 28/15 \end{pmatrix} = \text{upper triangular}$$

swap row 1 & 3 clear out col 1 clear out col 2

Then - detA = 2*(-15)*(28/15) = - 2*28 = - 56 ≠ 0.

```
m := matrix(3,3,[0,1,3,6,0,4,2,5,7]);
```

$$m := \begin{bmatrix} 0 & 1 & 3 \\ 6 & 0 & 4 \\ 2 & 5 & 7 \end{bmatrix}$$

```
det(m);
```

56

Proof #2. If the n columns of A are linearly independent, then column rank = n and by Theorem (A.3.1) the dimension of the range is also n. Since A: $E^n \rightarrow E^n$ for a square matrix, this range is all of E^n. According to Theorem (A.3.6), the dimension of the nullspace (the nullity) must then be n - n = 0, so A has only the trivial null space. We claim that the mapping A**x** = **y** is then one-to-one. Certainly each vector **x** goes into only one vector **y**. But could two domain vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ go into the same **y**? If they

did, then since the mapping is linear $\mathbf{x_1}$-$\mathbf{x_2}$ maps into 0. But this is a contradiction since we just showed that the nullspace can only contain the vector 0, whereas $\mathbf{x_1}$-$\mathbf{x_2} \neq 0$ if $\mathbf{x_1}$ and $\mathbf{x_2}$ are different. Thus, $\mathbf{x_1}$ and $\mathbf{x_2}$ must be the same, and we then have a 1-to-1 transformation. That fact means that the transformation must be invertible. We know that $A^{-1} = \text{cof}(A^T)/\det(A)$, so for A^{-1} to exist we must have $\det(A) \neq 0$.

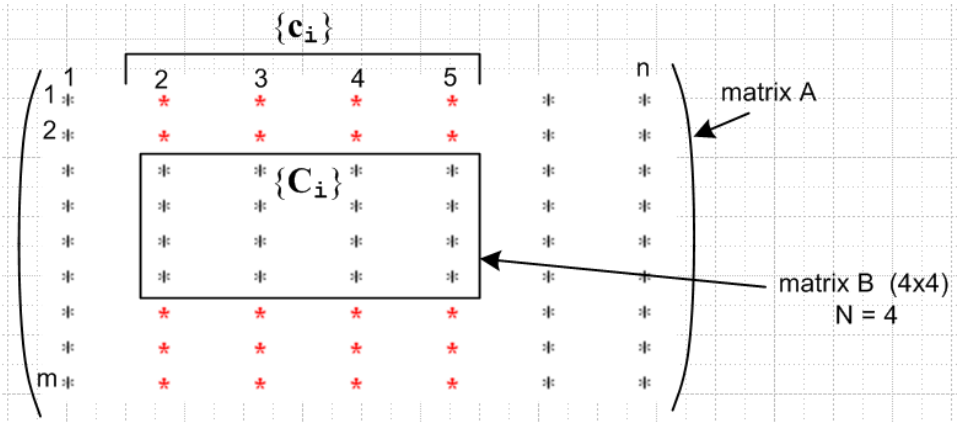
The following theorem involves a matrix A and a submatrix B.

Theorem. Consider a matrix with m rows and n columns. We can write A = ($\mathbf{c_1}$, $\mathbf{c_2}$....$\mathbf{c_n}$). Pick a subset $\{\mathbf{c_i}\}$ of these columns which contains N column vectors where N ≤ n and N ≤ m. Then create N reduced-size column vectors $\{\mathbf{C_i}\}$ by knocking out some set of components from all these $\{\mathbf{c_i}\}$ such that the resulting $\mathbf{C_i}$ vectors have exactly N components. For example, we might remove rows 2, 3 and 5 of all the $\{\mathbf{c_i}\}$ to create the $\{\mathbf{C_i}\}$ Construct an NxN matrix B from these N $\mathbf{C_i}$ vectors. This matrix B is of course some NxN submatrix of the larger nxm matrix A. Here then is the theorem: If the $\{\mathbf{C_i}\}$ form a set of linearly independent vectors within the NxN matrix B, then the set $\{\mathbf{c_i}\}$ form a set of linearly independent vectors for the larger matrix A.

theorem: $\{\mathbf{C_i}\}$ linearly independent in B $\Rightarrow$ $\{\mathbf{c_i}\}$ linearly independent in A

contrapositive: $\{\mathbf{c_i}\}$ linearly dependent in A $\Rightarrow$ $\{\mathbf{C_i}\}$ linearly dependent in B (A.4.3)

Before stating the trivial proof, here is a picture to clarify the setup for this theorem:



(A.4.4)

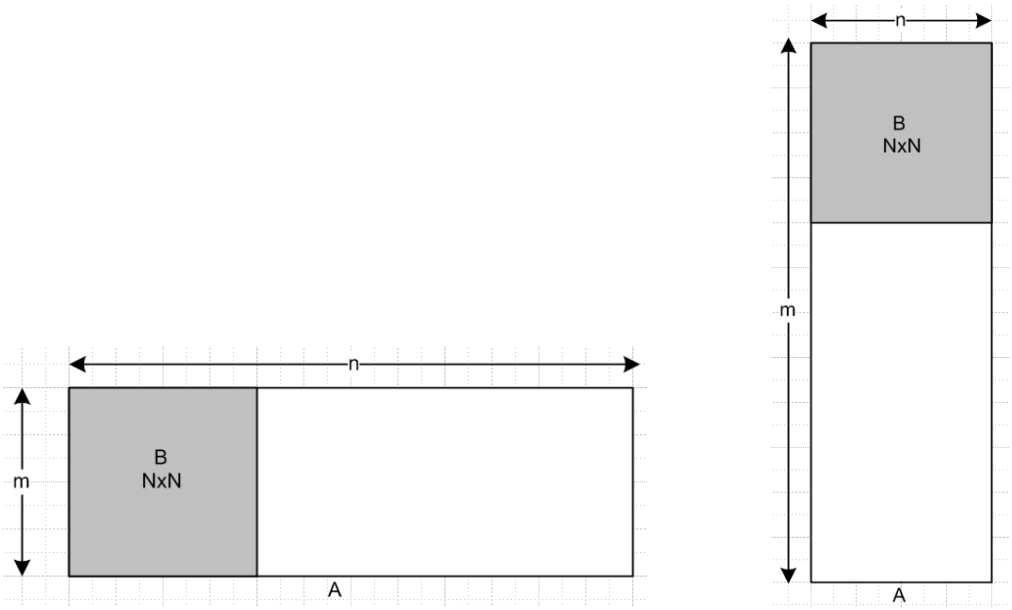
We happened to select for $\{\mathbf{c_i}\}$ the contiguous set of columns $\{\mathbf{c_2}, \mathbf{c_3}, \mathbf{c_4}, \mathbf{c_5}\}$. This causes the 4x4 matrix B to be contiguous in this picture. If the selected column set $\{\mathbf{c_i}\}$ were non-contiguous, one creates the set $\{\mathbf{C_i}\}$ and assembles them into a matrix B off to the side.

Proof of the contrapositive in the context of Fig (A.4.4) : If the set $\{\mathbf{c_i}\}$ is linearly dependent, then one can write $\Sigma_{i=2}^{5}\, k_i\mathbf{c_i} = 0$. This is a set of m equations which contains as a subset the set of N equations $\Sigma_{i=2}^{5}\, k_i\mathbf{C_i} = 0$. This last equation then says that the $\{\mathbf{C_i}\}$ are linearly dependent. QED.

Theorem. Consider a matrix A with m rows and n column. This matrix has some column rank N which lies in the range $0 \le N \le N_{\mathbf{max}} \equiv \min(n,m)$. Then the largest non-vanishing subdeterminant (minor) of A is of size NxN.

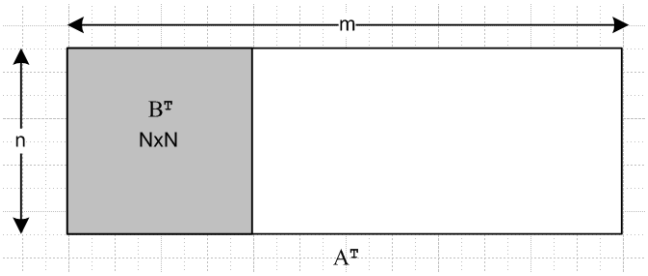
column rank = N $\Rightarrow$ largest non-vanishing subdeterminants are NxN (A.4.5)

Case 1. (full rank). Consider first a left shape for matrix A, where N = m:



We *select* for the NxN matrix B a matrix which contains the N linearly independent columns. This matrix B is likely to be disjoint, but we schematically draw it as a continuous matrix. Since the columns of B are linearly independent, we know from Theorem (A.4.2) that $\det(B) \neq 0$. Since B is an NxN submatrix of A, we have fulfilled the claim of the theorem.

On the right, we do NOT know that the columns of B are linearly independent. We only know that the total tall columns have this property! So now I need my dreaded theorem: if *all* det B = 0, then the columns cannot be linearly independent. What happens if I transpose the picture on the right? Get this,



I now know that the n rows of A^T are line indep, and this does not solve the problem. I need that theorem!

Suppose N = n, so the entire column set is a linearly independent set. Think of these as "tall" vectors. Consider any nxn square submatrix B within A. Such a square matrix B has n independent columns (why!) and this has full column rank. We know from Theorem (A.4.2) that det(B) ≠ 0. In this case, then, the largest non-vanishing subdeterminant with A has dimension n x n. This agrees with the theorem.

Case 2. Suppose N = n - 1. Imagine that the N independent columns are the leftmost ones. Consider any (n-1)x(n-1) square matrix B located within these columns within A. Such a matrix B has full rank, and for any such B we know from Theorem (A.4.2) that detB ≠ 0.

Now consider any nxn matrix B within A. The full column set which passes through such a B matrix is linearly dependent. According to Theorem (A.4.3), the smaller columns of B are also linearly dependent. Therefore this matrix B has less than full column rank so according to Theorem (A.4.1) det B = 0. Thus all nxn matrices within A have detB = 0, whereas there are many (n-1)x(n-1) submatrices which have non-zero determinant. In this case the largest non-vanishing subdeterminants are (n-1)x(n-1), and this agrees with our theorem claim.

Case 3. Suppose N = n - 2. Imagine that the N independent columns are the leftmost ones. Consider any (n-2)x(n-2) square matrix B located within these columns within A. Such a matrix B has full rank, and for any such B we know from Theorem (A.4.2) that detB ≠ 0.

(n-1)x(n-1) square matrices B. Since there are only n-2 independent full columns in A, any full column set containing an (n-1)x(n-1) square matrices B is linearly dependent. By theorem (A.4.3) the columns within B are also dependent, and thus detB = 0.

What about n x n square matrices? These are embedded in a set of linearly dependent columns, so these all have detB = 0.

I think I have it, but presentation needs to be cleaned up. The upshot is that if column rank = N, then the largest det(B) ≠ 0 submatrices are of dimension N x N.

Now I want the converse of the above theorem, so

Theorem. largest non-vanishing subdeterminants are NxN $\Rightarrow$ column rank = N (A.4.6)

Case 1. Suppose N = n. This means there is at least one nxn matrix B which has detB ≠ 0. Within such a matrix B the columns are independent. According to Theorem (A.4.3) the full columns are also independent and therefore column rank = N.

Case 2. Suppose N = n-1. This first of all means that det B = 0 for any nxn submatrix. This means the that columns C_i in any such B matrix are dependent. But now (A.4.3) is of no help! I therefore arrive at no conclusion for the full columns c_i being dependent or independent.

I do know that there is at least one (n-1)x(n-1) matrix B which has detB ≠ 0. (A.4.3) then tells us that there are n-1 full columns which are independent. If all n full columns were independent, then we would know that all nxn matrices have detB ≠ 0. But this is a contradiction, so it must be that n-1 columns are independent and so column rank is N = n-1, and we are good.

Case 2. Suppose N = n-2. We know at once that all nxn and (n-1)x(n-1) dets are 0 and at least one $(n-2)^2$ det is non 0. That last fact says that n-2 full columns are lin indep. If n-1 were, then etc etc.

I think I have it. I then never need this theorem:

Not needed: If all det B = 0 for a set of columns, then the full columns are dependent.

Theorem: If the largest non-vanishing subdeterminant in general matrix A has size NxN, then the largest non-vanishing subdeterminant in A^T also has size NxN. (A.4.7)

Proof. Suppose general matrix A has a square submatrix B. Then the matrix A^T has the square submatrix B^T. Note that det(B) = det(B^T). If all square B within A of size larger than NxN have det(B) = 0, then it follows that all square B within A^T of size larger than NxN also have det(B) = 0. This is to because B within A^T implies B^T within A. But B^T in A is just some square matrix B' in A and if B' is larger than NxN, then det (B') = 0. Since det(B') = det(B^T) = det(B), we get det(B) = 0. Similarly, if there exists a square B within A of size NxN which has det(B) $\neq$ 0, then it follows that there exists a square B within A^T of size NxN which has det(B) $\neq$ 0.

Theorem. For any mxn matrix A, column rank = row rank = rank. (A.4.8)

Proof: Suppose column rank(A) = N $\leq$ min(n,m). Then by Theorem (A.4.5) the largest non-vanishing subdeterminants of A are NxN. By Theorem (A.4.7) the largest non-vanishing subdeterminants of A^T are also of size NxN. From Theorem (A.4.6), the column rank of A^T is N. But column rank(A^T) = row rank(A), and therefore row rank(A) = N. Thus, the column rank and row rank of matrix A are both N, so they are equal.

We may now restate certain earlier theorems using "rank" in place of "column rank" :

Theorem. In the linear transformation context A : $E^n \rightarrow E^m$, the dimension of the range of the transformation is equal to the rank of the mxn matrix A.

dim(R) = rank(A) (A.3.1)'

Theorem. Consider a matrix A with m rows and n columns. Then

nullity(A) + rank(A) = n. OR dim($\mathcal{N}$) + dim(R) = n (A.3.6)'

Theorem. For an n x n square matrix A, if det(A) $\neq$ 0, then the rank is n. (A.4.1)'

Contrapositive: rank < n $\Rightarrow$ det(A) = 0.

Theorem. For an n x n square matrix A, if the rank is n, then det(A) $\neq$ 0. (A.4.2)'

Theorem. Consider a matrix A with m rows and n columns. This matrix has some rank N which lies in the range 0 $\leq$ N $\leq$ min(n.m). Then the largest non-vanishing subdeterminant of A has size NxN.

rank = N $\Rightarrow$ largest non-vanishing subdeterminants are NxN (A.4.5)'

Theorem. largest non-vanishing subdeterminants are NxN $\Rightarrow$ rank = N (A.4.6)'