Notes on Rank PhL 5.29.15

On page 231 the Bucks make these statements:

Def 1. "The rank of any matrix, square or not, is the size of the largest square submatrix which can be obtained by erasing rows and columns and which has a non-zero determinant. "

Def 2. "The rank can also be defined as the largest number of rows (or columns) of the matrix which form an independent set. "

Def 3. "The rank is the dimension of the range of the transformation implied by the matrix. "

The goal of this little document is to show why the above three definitions of rank are all the same. Right now, the three definitions seem to me to be totally unconnected. This is partly because my matrix knowledge has evaporated over the last several years due to disuse, just as piano skill evaporates without practice.

Part A. Comparing Def 2 and Def 3.

A.1. Start with Square Matrices and Definition 2.

(a) Suppose the three columns of a 3x3 matrix are linearly independent. The matrix is A = $(\mathbf{c}_1\mathbf{c}_2\ \mathbf{c}_3)$. If I apply the matrix A to the three unit vectors like (1,0,0) which span the domain, what happens?

$$(\mathbf{c}_1\ \mathbf{c}_2\ \mathbf{c}_3)\begin{pmatrix}1\\0\\0\end{pmatrix} = \mathbf{c}_1 \qquad (\mathbf{c}_1\ \mathbf{c}_2\ \mathbf{c}_3)\begin{pmatrix}0\\1\\0\end{pmatrix} = \mathbf{c}_2 \qquad (\mathbf{c}_1\ \mathbf{c}_2\ \mathbf{c}_3)\begin{pmatrix}0\\0\\1\end{pmatrix} = \mathbf{c}_3$$

I obtain in the range the three vectors $\mathbf{c}_i$ as shown. But I assumed these were linearly independent, so the range must have dimension 3. One might write this in a single shot

$$(\mathbf{c}_1\ \mathbf{c}_2\ \mathbf{c}_3)\begin{pmatrix}\alpha\\\beta\\\gamma\end{pmatrix} = \alpha\,\mathbf{c}_1 + \beta\,\mathbf{c}_2 + \gamma\,\mathbf{c}_3 = \text{an arbitrary vector in the range.}$$

(b) Suppose only two of the three columns are linearly independent. Suppose specifically that $\mathbf{c}_1$ and $\mathbf{c}_2$ are linearly independent, and therefore we can write $\mathbf{c}_3 = k_1\mathbf{c}_1 + k_2\mathbf{c}_2$ so it is "dependent". In this case we find that

$$(\mathbf{c}_1\ \mathbf{c}_2\ \mathbf{c}_3)\begin{pmatrix}\alpha\\\beta\\\gamma\end{pmatrix} = \alpha\,\mathbf{c}_1 + \beta\,\mathbf{c}_2 + \gamma\,\mathbf{c}_3 = \alpha\,\mathbf{c}_1 + \beta\,\mathbf{c}_2 + \gamma\,[k_1\mathbf{c}_1 + k_2\mathbf{c}_2\ ]$$

$$= (\alpha + \gamma k_1)\mathbf{c}_1 + (\beta + \gamma k_2)\mathbf{c}_2 = A\mathbf{c}_1 + B\mathbf{c}_2$$

Now the range is spanned by only two vectors $\mathbf{c_1}$ and $\mathbf{c_2}$ and this the range has dimension 2. The vectors are not collinear because we have said they are independent. Our conclusion here would be the same no matter which pair of c_i we assumed were linearly independent -- the range has dimension 2.

(c) Suppose only one column is linearly independent, perhaps c_1. Then you can write $\mathbf{c_2} = k_2\mathbf{c_1}$ and $\mathbf{c_3} = k_3\mathbf{c_1}$. Doing the above, we then end up with

$$(\mathbf{c_1}\ \mathbf{c_2}\ \mathbf{c_3}) \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = A\mathbf{c_1}$$

and we conclude that the range has dimension 1.

Conclusion: I think I could make this general, and this seems to glue together the two notions of rank stated in Def 2 and Def 3 above, at least for square matrices.

A.2. Generalize Section 1 to non-square matrices where #rows > #cols.

(a) Let's keep the domain of dimension 3. But now allow each column $\mathbf{c_i}$ to have 4 elements.

$$(\mathbf{c_1}\ \mathbf{c_2}\ \mathbf{c_3}) \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \alpha\,\mathbf{c_1} + \beta\,\mathbf{c_2} + \gamma\,\mathbf{c_3}$$

The mapping here is $A : E^3 \rightarrow E^4$. The range *space* has dimension 4. Suppose the three $\mathbf{c_i}$ are linearly independent in E^4 (where they live). You can see then that the range has dimension 3, despite the fact that the range space has dimension 4. This is because the range is spanned by 3 vectors as shown.

(b) If only two of the c_i are linearly independent, we will get dim(range) = 2, exactly as in the previous section.

Conclusion: For non-square matrices with #rows > #cols, the conclusions are exactly the same as for the case of a square matrix, but I guess we have to speak of the "column rank" being the number of independent column vectors. So column rank = dimension of the range. The "row rank" does not enter the discussion at all. There could be 20 rows, and each has 3 entries, and perhaps the row rank = 3 all the time, despite that column rank varying. We only care about the column rank because the column count is the smaller of the two counts.

A.3. Generalize Section 1 to non-square matrices where #rows < #cols.

(a) Again keep the domain dimension at 3. Again write

$$(\mathbf{c_1}\ \mathbf{c_2}\ \mathbf{c_3}) \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \alpha\,\mathbf{c_1} + \beta\,\mathbf{c_2} + \gamma\,\mathbf{c_3}$$

but assume there are only 2 rows, The mapping here is A: $E^3 \rightarrow E^2$. Each $\mathbf{c_i}$ vector has 2 components only. In this case, the three $\mathbf{c_i}$ cannot be linearly independent. The dimensionality of the range is pinned at 2 which is the dimension of the range space. The Def 2 rank of the matrix A can be at most 2 because we could have two of the $\mathbf{c_i}$ being linearly independent. The column rank here cannot exceed 2 and this column rank matches the dimension of the range.

(b) Consider the more general case A : $E^N \rightarrow E^n$ where N > n. The matrix has n rows and N columns. You can have at most n of the $\mathbf{c_i}$ being linearly independent, so rank ≤n by Def 2. Let's write this schematically,

$(\mathbf{c_1}\ \mathbf{c_2}\ \mathbf{c_3}\ .....\mathbf{c_N})\ \mathbf{v} = v_1\ \mathbf{c_1} + v_2\ \mathbf{c_2} + v_3\ \mathbf{c_3}\ .... v_N\ \mathbf{c_N}$ = dimension range of n.

Suppose the first n of the c_i are lin indep. Then you can write all the remaining c_i as lin com of the first once, and you end up with $A\mathbf{c_1} + B\mathbf{c_2} + ... Q\mathbf{c_n}$ on the right, and you are pinned at dimension n.

Conclusion. In this case we can say the following:

- the max number of independent columns $\mathbf{c_i}$ is n, so column rank ≤ n.
- the dimension of the range is the column rank of A.

A.4. Try to make a statement covering all three cases.

Assume there are n rows and m columns.

	rows	cols		max col rank	dim range
square	n	m	n=m	n	col rank
tall	n	m	n>m	m	col rank
wide	n	m	m>n	n	col rank

So I think I have now established that Def 2 and Def 3 agree for any matrix.

Part B. Comparing Def 2 and Def 1.

B.1 Assume a square matrix.

Suppose all n columns are lin indep. Then I know that detA ≠ 0. This is the content of my Theorem 9.5 which I state and prove on page 10 of Section 2 of the Matrix Binder. The proof is an algorithm that I invented for this purpose, but probably other proofs exist.

Suppose only 1 column is lin dep on the others. We know that in this case detA = 0. One way to know this is to do det-preserving operations to clean out an entire column and it is then all zeros.

Suppose only 2 columns are lin dep on the others. Any n-1 x n-1 submatrix you consider must include at least one of these columns. Any such submatrix therefore has det = 0 since within its reduced dimension it has a lin dep column.

Suppose only 3 columns are lin dep on the others. Any n-2 x n-2 submatrix you consider must include at least one of these columns. Any such n-2 x n-2 submatrix therefore has det = 0 since within its reduced dimension it has a lin dep column.

The general idea seems to be this:

If an nxn matrix has k+1 lin dep columns, then all n-k x n-k subdeterminants must vanish.

If an nxn matrix has k+1 lin dep columns, it must have n - (k+1) lin indep columns, and therefore it has a Def 2 rank of n-(k+1) = n - k -1. We have just shown that in this case, all n-k x n-k subdeterminants must vanish. Lets call the Def 2 rank by name r. Then r = n-k-1 $\Rightarrow$ n-k x n-k dets all vanish. Then

Def 2 rank = r $\quad\Rightarrow\quad$ r+1 x r+1 dets all vanish

Test this. If matrix has full rank, then r = n and all n+1 x n+1 dets all vanish, but there are no such dets and so we just have detA $\rightarrow$ 0.

If rank is one down, then r = n-1 and then all n x n dets vanish. There is only one such det.

If rank is two down, then r = n-2 and then all n-1 x n-1 dets vanish.

I think I am onto it now, but much nailing down must be done, all pieces are floating around right now.

****

B.1 Square matrices. Go back to our 3x3 case.

$$(\mathbf{c}_1\ \mathbf{c}_2\ \mathbf{c}_3)\begin{pmatrix}\alpha\\ \beta\\ \gamma\end{pmatrix} = \alpha\,\mathbf{c}_1 + \beta\,\mathbf{c}_2 + \gamma\,\mathbf{c}_3 = \text{an arbitrary vector in the range.}$$

What is the connection to "determinants" ? I know that (for example)

$$\det(A) = \varepsilon_{ijk}A_{1i}A_{2j}A_{3k}$$

Suppose the three columns are lin indep. How do we know that det A $\neq$ 0 ? This is the content of my Theorem 9.5 which I state and prove on page 10 of Section 2 of the Matrix Binder. The proof is an algorithm that I invented for this purpose, but probably other proofs exist.

Now suppose two columns $\mathbf{c_1}$ and $\mathbf{c_2}$ are linearly independent only. I know that detA = 0 in this case, But I don't know where to go now.

Look at Matrix Binder Section 1 page 9 where I give the Def 1 definition of "rank" and then I prove some things based on this definition, like Theorem R1. The details for this stuff are on page 7 of Section 4!