**Assigned Problems:**  1, 2, 3, 4, 7, 8, 12, 16, 18

This section contains additional results about the rank of a combination of matrices.

Let's start with a result we can easily prove:

\[
\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}
\]  

(1)

This gives a quick upper bound for rank(AB). It says that matrix multiplication cannot increase rank. From the column-wise view of matrix multiplication, we get \(\text{ColSp}(AB) \subseteq \text{ColSp}(A)\). From the row-wise view of matrix multiplication, we get \(\text{RowSp}(AB) \subseteq \text{RowSp}(B)\). From the first inclusion it follows that \(\text{rank}(AB) \leq \text{rank}(A)\) (why?), and from the second inclusion it follows that \(\text{rank}(AB) \leq \text{rank}(B)\) (why?). Hence \(\text{rank}(AB)\) is less than or equal to the smaller of \(\text{rank}(A)\) and \(\text{rank}(B)\), which is the result we wanted to prove.

The key result of Section 4.5 is a refinement of (1) that gives the exact value for \(\text{rank}(AB)\):

\[
\text{rank}(AB) = \text{rank}(B) - \dim (N(A) \cap R(B))
\]  

(2)

From this equation, many other useful results about the rank of a product can be derived, including (1). We will discuss the meaning of (2) and outline its proof. You should understand why (2) makes sense — why \(N(A)\) is involved, not \(N(B)\) or some other thing.

Using (2) you should know how to derive a lower bound for \(\text{rank}(AB)\):

\[
\text{rank}(A) + \text{rank}(B) - n \leq \text{rank}(AB)
\]  

(3)

Other topics you should pay attention to

- If \(\text{rank}(A) = r\), then there is at least one \(r \times r\) submatrix of \(A\) that is non-singular. And every larger submatrix of \(A\) will be singular.

- Small perturbations cannot reduce rank. The proof, when read carefully, shows that a small perturbation is likely to increase rank. Read “A Pitfall in Solving Singular Systems” (p. 217-8).

- Know the different ways of characterizing rank (p. 218).

Finally, I want to mention that this section also contains a discussion of the products \(A^T A\) and \(AA^T\). These products arise in the context of least squares problems. We will not dwell on this topic now. Be aware of the following:

- Suppose \(Ax = b\) is an inconsistent system. Here \(A\) is \(m \times n\), with typically \(m\) much larger than \(n\) (so more equations (i.e. constraints) than variables). The new system formed by multiplying on the left by \(A^T\), that is the system \(A^T Ax = A^T b\) is nevertheless always consistent, and its solution is the least squares solution to the original system \(Ax = b\).

- However. This is NOT the way least squares solutions are found in practice. Calculating the matrix product \(A^T A\) magnifies any sensitivity the original system \(Ax = b\) may have had to round-off error (see Example 4.5.1, p. 214-5).

I will not cover section 4.6 in class — it deals with how to solve least squares via the normal equations \(A^T Ax = A^T b\). If the semester were longer, we could linger over methods like these but... The preferred way to solve least squares is to use the QR-factorization of \(A\) — which we will study in Chapter 5. I do, however, urge you to read the Epilogue on p. 233-4, which is a wonderful historical note on Gauss, and the circumstances that led him to invent the theory of least squares.