HW7

19.1 Given $A \in \mathbb{C}^{m \times n}$ of rank n and $b \in \mathbb{C}^m$, consider the 2×2 block system of equations

$$\begin{bmatrix} I & A \\ A^* & 0 \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}$$

where $I \in \mathbb{R}^{m \times m}$. Show that this system has a unique solution $(r, x)^T$ and that the vectors r & x are the residual and the solution of the least squares problem (18.1)

$$\begin{pmatrix} I & A \\ A^* & 0 \end{pmatrix} \begin{pmatrix} r \\ x \end{pmatrix} = \begin{pmatrix} r + Ax \\ A^* r \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$ found by multiplication

In other words, $r = b - Ax$ and $A^* r = 0$. These are equations (11.1) and (11.8) which according to Theorem 11.1 are equivalent to solving least squares problem

19.2 Interpret code:

```
[U,S,V] = svd(A);
S = diag(S);
tol = max(size(A))*S(1)*eps;
r = sum(S>tol);
S = diag(ones(r,1)./S(1:r));
X = V(:,1:r)*S*U(:,1:r)';
```

The least squares problem is not well defined when A is rank deficient UNLESS you add the condition that $\|x\|$ is minimized. The matrix X that is considered above when applied to b will produce the x's that solves the rank deficient least squares problem. The code deals with the problem that rank deficiency is lost when numerical calculations are made. It replaces the test for zero singular values with a test for singular values that are as small as could be expected to be detected with the svd calculation. That calculation produces errors that are of a size proportional to $\|A\|\varepsilon_m$. The largest singular value is $\|A\|$. The proportionality constant is taken to be maximum of the dimensions of A. Certainly we can expect that the numerical error should grow with the size of the matrix A, though it is hard to see exactly what the factor should be.

To see why Xb is a solution, assume the near zero singular values really were zero, and note that $b - AXb$ is perpendicular to the first r left singular vectors (a basis for the range of A). Why? because $AXb = U S V^* \hat{V} \hat{S} \hat{U}^* b$ where $\hat{V} = V(:, 1:r)$, $\hat{U} = U(:, 1:r)$ and $\hat{S}$ is the diagonal matrix whose entries are the inverses of the first r singular values. As seen before the zeros in the S matrix mean that the rows beyond row r make no contribution.

$$A X b = \hat{U} \hat{S} \hat{V}^* \hat{V} \hat{S} \hat{U}^* b = \hat{U} \hat{S} \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \hat{S} \hat{U}^* b = \hat{U} \hat{U}^* b$$

This is the projection of b onto the span of the columns of U. As we saw in Lecture 11, any other solution minimizing the residual will differ by something in the nullspace of A, that is, by something perpendicular to the columns of $\hat{V}$. To put this in another way, by a linear combination of the $(r+1)^{st}$ through n^{th} columns of V. By construction, Xb is a linear combination of the first r columns so adding something perpendicular can only increase the norm (Pythagoras' Theorem). This shows that Xb minimizes the residual and $\|x\|$. This is the right generalization of the pseudo-inverse.

22.1 Show that for Gaussian elimination with partial pivoting applied to any matrix $A \in \mathbb{C}^{m \times m}$ the growth factor (22.2) satisfies $\rho \leq 2^{m-1}$

Suffices to show that for each $m-1$ steps of the elimination algorithm the ratio of the maxima of the entries of the "after" A^{j+1} and "before" A^j matrices is at most 2. Exchanging rows will not change the value of the maximum, so we may assume that the pivot row (j^{th} row) is already in place so that the pivot has maximum absolute value for entries in its column in the lower triangle. For $k > j$ the entries A^{j+1} are

$$a_{ki}^{j+1} = a_{ki}^{j} - \frac{a_{kj}^{j}}{a_{jj}^{j}} a_{ji}^{j}$$

while the other entries are unchanged. Taking the absolute value

$$|a_{ki}^{j+1}| \leq |a_{ki}^{j}| + \frac{|a_{kj}^{j}|}{|a_{jj}^{j}|} |a_{ji}^{j}|$$

But pivots are such that $\frac{|a_{kj}^{j}|}{|a_{jj}^{j}|} \leq 1$

$$\therefore \quad |a_{ki}^{j+1}| \le |a_{ki}^{j}| + |a_{ji}^{j}| \le 2 \max_{1 \le p,q \le m} |a_{pq}^{j}|$$

Thus the maximum entry is no worse than the factor of 2 larger than that of "before" matrix.

21.6 Suppose $A \in \mathbb{C}^{m \times m}$ is "strictly column diagonally dominant", which means that for each k

$$|a_{kk}| > \sum_{j \neq k} |a_{jk}|$$

Show that Gaussian elimination with pivoting does not require row exchanges (i.e. possibly only scaling).

We need to show that $(m-j+1 \times m-j+1)$ lower left block of the A_j produced via Gaussian elimination are all diagonally dominant if the starting matrix $A_1 = A$ is diagonally dominant. Without loss of generality we can look at a single step, for example, the first step.

Since $|a_{11}| > \sum_{j \neq 1} |a_{j1}| \ge |a_{j1}|$ we know that a_{11} is the pivot (no row exchange needed).

Let $B = A_2 = L_1 A_1 = L_1 A$.

Then $b_{kj} = a_{kj} - \frac{a_{k1}}{a_{11}} a_{1j}$. So

$$|b_{kj}| = \left|a_{kj} - \frac{a_{k1}}{a_{11}} a_{1j}\right| \leq |a_{kj}| + \frac{|a_{k1} a_{1j}|}{|a_{11}|}$$

Summing:

$$\sum_{k \neq 1,j} |b_{kj}| \leq \sum_{k \neq 1,j} |a_{kj}| + \frac{|a_{1j}||a_{k1}|}{|a_{11}|}$$

$$= \sum_{k \neq j} |a_{kj}| - |a_{1j}| + \frac{|a_{1j}|}{|a_{11}|} \sum |a_{k1}| - \frac{|a_{1j}||a_{j1}|}{|a_{11}|}$$

However the diagonal dominance of Δ applied to column j and to column 1 allows us to bound the summation terms by the corresponding diagonal terms

$$\sum_{k \neq 1,j} |b_{kj}| \leq |a_{jj}| - |a_{1j}| + \frac{|a_{1j}|}{|a_{11}|} |a_{11}| - \frac{|a_{1j}|}{|a_{11}|} |a_{j1}|$$

$$= |a_{jj}| - \frac{|a_{1j}|}{|a_{11}|} |a_{j1}|$$

But $b_{jj} = a_{jj} - \frac{a_{j1}}{a_{11}} a_{1j}$ we know that

$|b_{jj}| \geq |a_{jj}| - \frac{|a_{j1}|}{|a_{11}|} |a_{1j}|$. Hence

$$\sum_{k \neq j} |b_{kj}| \leq |b_{jj}|$$

That is, that the lower $m-1 \times m-1$ block digonal dominant, which is what we set out to prove.

22.3 The Growth factor vs dimension shows that, when Gaussian elimination with partial pivoting is applied to matrices whose entries are chosen randomly from a uniform distribution on $[-1,1]$, these fall <u>below</u> the $n^{3/2}$ line.
The $n^{1/2}$ & $n^{3/2}$ lines are superimposed.

This is different from what appears in the text, which corresponds to matrices with entries picked from a normal distribution.

The pdf plot shows the approximate pdf for the growth factor for $n = 8, 16, 32$. In each case the tails decay exponentially.
The slopes/peaks reflect the fact that we're using a uniform distribution instead.

The next 4 figures refer to Trefethen's discussion on L^{-1} entries and $\hat{L}$ (see text)
The upshot is that similar results are found when the normal distribution is replaced by a uniform distribution