CS 621: Matrix Computations
Fall 2001
Problem Set 2

Handed out: Fri., Sep. 21.

Due: Mon., Oct. 1 in lecture.

1. The inverse of an $n \times n$ nonsingular lower triangular matrix can be found, one column at a time, by repeatedly applying forward substitution to systems of the form $L\mathbf{x} = \mathbf{e}_i$, where $\mathbf{e}_i$ denotes the ith column of the identity matrix. How many flops, accurate to the leading term, are required for this computation? Note that many flops can be saved because you can determine a priori that many entries of the inverse are zeros.

2. (GVL P2.2.3) Let $D \in \mathbf{R}^{n\times n}$ be diagonal. Show that $\|D\|_p = \max_i |D(i,i)|$ for any $p \in [1,\infty]$. [Hint: prove separately that $\|D\|_p \le \max_i |D(i,i)|$ and $\|D\|_p \ge \max_i |D(i,i)|$.]

3. Consider the "Sort" operation that takes as input a vector in $\mathbf{R}^n$ and returns the entries in sorted (least-to-greatest) order. Show that this operation is well-conditioned for all inputs. In other words, prove that a small normwise relative change to the input vector leads to a small normwise relative change to the output vector.

   [Hint: To simplify notation, renumber the subscripting of $\mathbf{x}$ according to the order in Sort($\mathbf{x}$), so that $x(1) \le x(2) \le \cdots \le x(n)$. The easier case is when a particular perturbed $x(i) + \delta(i)$ ends up also at position i in the sorted perturbed list. The harder case is when $x(i) + \delta(i)$ ends up at a different position $j \ne i$. Start with the case $j < i$: Argue that on the one hand, $x(i) + \delta(i)$ cannot be much smaller than $x(j)$ since $x(j) \le x(i)$. Argue on the other hand that $x(i)+\delta(i)$ cannot be much larger than $x(j)$ because otherwise there would be at least j entries in the sorted perturbed list ranked smaller than $x(i) + \delta(i)$ (why?), contradicting the assumption that $x(i) + \delta(i)$ landed at position j.]

4. Consider the polynomial $p(x) = (x - 1.1)^k$. Write a script in Matlab that plots this function (by evaluating it at closely spaced points) for the domain $[.5, 1.5]$. Then multiply the factors of $p(x)$ together to obtain the array of coefficients in the standard-form representation $a_0 + a_1x + \cdots + a_kx^k$ of $p(x)$ and evaluate the expanded form of the polynomial again over the interval $[.5, 1.5]$ and plot this. (Use `poly` to obtain the coefficients and `polyval` to evaluate the polynomial.)

   You should notice a substantial difference between the two plots as k increases. The reason for the difference is that the second plot suffers from cancellation error. Why? (Hint: look at the coefficients in the expanded form. If a sequence of numbers with large absolute values are added together to yield a number with small absolute value, then the answer will suffer from cancellation.)

   The largest coefficient of the expanded polynomial will be close to the middle. (Assume k is even.) This coefficient is derived from the binomial formula, which in turn involves factorials. Using Stirling's approximation for factorials to come up with an inequality, in terms of k and u (where u is unit roundoff), for what is the largest value of k before the relative error due to cancellation reaches 100%. (If you are not familiar with Stirling's approximation, you can locate it on the web.)

Hand in: your m-file, the requested plots, and the analysis described in the last paragraph.