Prof. Ming Gu, 861 Evans, tel: 2-3145
Office Hours: TuWTh 12:00-1:30PM
Email: mgu@math.berkeley.edu
http://www.math.berkeley.edu/∼mgu/MA2212008F

Math221: Matrix Computations

Selected Solutions to Homework #1, 2008

- Problem 1.9: We will only consider the computations with tiny $x > 0$. Let ϵ be the machine precision and let $x = (k + f)\epsilon$, where k is an integer satisfying $10 > k \geq 0$ and f is the fraction with $1 > f \geq 0$. In all cases, we need to compute $d = 1 + x$. We assume that $\mathbf{fl}(d)$ is obtained via rounding to the nearest. This implies that $\mathbf{fl}(d) = 1 + k\epsilon$ if $f < 1/2$ and $\mathbf{fl}(d) = 1 + (k + 1)\epsilon$ otherwise. This discontinuity in floating point computation of $1 + x$ is the cause of the large errors in left plots. On the right plots, since $d = 1 + x$ has already been computed and is stored as $\mathbf{fl}(d)$, both $\log(\mathbf{fl}(d))$ and $\mathbf{fl}(d) - 1$ can be computed highly accurately with respect to $\mathbf{fl}(d)$.
- **overflow/underflow protection in quadroot.m:** On the class website, there is a modified version of quadroot.m called quadroot2.m as well as a paper on solving for quadratic roots by George Forsythe.