Ed Bueler April 25, 2013

# Review of Turing (1948)

Full citation: A. M. Turing (1948). "Rounding-off errors in matrix processes," *The Quarterly Journal of Mechanics and Applied Mathematics*, **1** (1), 287–308.

My summary[1]: Turing starts numerical linear algebra from the beginning. His major concern is the disaster-predicting claim from Hotelling (1943) that "in solving a set of $n$ equations we should keep $0.6n$ extra" decimal digits. (So 600 digit arithmetic is needed to solve 1000 linear equations!)

He tries to establish that such exponential growth of error does not generally occur when solving systems of equations. He introduces the condition number of a matrix for this. He attempts a forward error analysis of Gauss elimination with partial pivoting, giving a (not proven) estimate that the error is polynomial in $m$, instead of exponential. He does not use (i.e. discover) the backwards-error/stability analysis of Wilkinson (1961). He provides example matrix (22.4). His last sentences on this topic predict the 1990s work on the errors in Gauss elimination.

At the same time, von Neumann and others thought one should use the normal equations and Cholesky (Alg 23.1). Turing explains why not: it makes the problem more ill-conditioned.

Relation to topics in TREFETHEN & BAU: Turing lays out almost all of Lecture 20 from the modern point of view, including "$J_{n-1} \dots J_1 A = DU$" as his version of equation (20.1). His first theorem is the solution to exercise 20.1: if the principal minors of $A$ are nonsingular then there is a unique LU factorization. For both Alg 20.1 and the partial pivoting version (Alg 21.1), he gives a modern matrix-factorization view. He writes the sequence $A = LU$, $Ly = b$, $Ux = y$ for how to use LU to solve systems.

He defines the "$N$-condition number" (compare Lecture 12) of $A$ as $n^{-1}\|A\|_F\|A^{-1}\|_F$. He has part of the idea of Theorem 12.2, but non-deterministic (and not proven). Specifically, he asserts that solving systems with random matrices makes errors $\delta x$ such that $\frac{\|\delta x\|_2}{\|x\|_2} = n^{-1}\|A\|_F\|A^{-1}\|_F \frac{\|\delta A\|_F}{\|A\|_F}$.

Also: He hints at the QR decomposition (Lecture 7) as "an upper triangular matrix $M$ such that $M^*A^*AM = I$, that is, so that $AM$ is orthogonal," but neither Gram-Schmidt or Householder is suggested. He defines the following matrix norms (Lecture 3): $\|A\|_F$, $\|A\|_2$, and the $\infty$-norm of $A$ treated as a vector; I think he knows $\|A\|_2$ is more fundamental but he does not use it after defining it. He introduces the term "preconditioning" (Lecture 40) for the first time, noting—this is 21st century stuff!—that it requires "considerable liaison between the experimenter and the computer," now called "physics-based preconditioning".

Most insightful/interesting/curious ideas: His way of counting operations includes "recordings of numbers, and extractions of figures from tables." He asserts (wrongly) that "with the advent of electronic computers it will become standard practice" to compute $A^{-1}$ to solve $Ax = b$, but then he knows this is at extra (three-times) cost. He describes LU as "$A = LDU$" factorization, where both $L$ and $U$ have unit diagonal and $D$ is diagonal, but then groups $DU$ as we do modernly. He states that determinant is *not* a good measure of conditioning, giving four $3 \times 3$ matrices with very different condition numbers and the same determinant. He conjectures that random matrices have $N$-condition number of $m^{1/2}$—proven much later—thus that "random matrices are only slightly ill-conditioned".

[1] Not considered here: He also analyses "Morris's escalator method"—never heard of it!—and Gauss-Jordan.