Numerical Linear Algebra

Sheet 3 — MT20

Stability and eigenvalues

This sheet is due 9am two weekdays before the class (Thursday if class is on Monday).

1. Stability of matrix multiplication. Denote by $fl(X)$ a computed approximation to X.

   It can be shown that vector-vector multiplication is backward stable, in that $fl(a^Tb) = (a+\Delta a)^Tb = a^T(b+\Delta b)$ for some small $\Delta a, \Delta b$: $\|\Delta a\|/\|a\| = \epsilon$ and $\|\Delta b\|/\|b\| = \epsilon$. Here the $\epsilon = O(u)$ notation supresses low-degree polynomials in m, n, that is, we write e.g. $umn = \epsilon$.

   (a) Use this to prove that matrix-matrix multiplication involving orthogonal matrices (is backward stable, that is,)←satisfies $\|fl(QA) - QA\| = \epsilon\|QA\|$ and $\|fl(AQ) - AQ\| = \epsilon\|AQ\|$.

   (b) Then prove that for any A, B (nonorthogonal, where A is square or tall and B is square or fat), $\|fl(AB) - AB\| \le \epsilon\|A\|\|B\|$ and hence obtain the general error bound for matrix multiplication

   $$\|fl(AB) - AB\|/\|AB\| \le \epsilon \min(\kappa_2(A), \kappa_2(B)).$$

   When is the upper bound sharp (up to the $\epsilon = O(u)$ notation), and when is it an overestimate?

2. (a) If $Ax = b$ and $(L+\Delta L)(U+\Delta U)\hat{x} = b$ with $LU = A+\Delta A$, find a δA that satisfies $(A+\delta A)\hat{x} = b$.

   (b) A computed pivoted LU factorisation $\hat{L}\hat{U} \approx PA$ is known to satisfy $\|\hat{L}\hat{U} - PA\| = \epsilon\|\hat{L}\|\|\hat{U}\|$. Use this to show that if $\|\hat{U}\| = O(\|A\|)$, then a linear system is solved in a backward stable manner, that is, the computed $\hat{x}$ satisfies (P)←$(remove)$$(A+\Delta A)\hat{x} = b$ where $\|\Delta A\| = \epsilon\|A\|$. (you can use the fact that triangular systems are backward stable).

Yuji Nakatsukasa: nakatsukasa@maths.ox.ac.uk

3. The (unshifted) QR algorithm performs: $A_1 = A$, and for $k = 1, 2, \ldots$ form the QR factorisation $A_k = Q_k R_k$, and set $A_{k+1} = R_k Q_k$.

   (a) Prove that
   $$A^k = (Q_1 \cdots Q_k)(R_k \cdots R_1) = Q^{(k)} R^{(k)} \tag{1}$$
   is the QR factorisation of A^k, and
   $$A_{k+1} = (Q^{(k)})^T A Q^{(k)}. \tag{2}$$

   (b) Suppose that $A \in \mathbb{R}^{n\times n}$ is symmetric. First show that
   $$A^{-k} (= (A^{-1})^k) = (R^{(k)})^{-1} (Q^{(k)})^T.$$
   Taking the transpose yields
   $$(A^{-k})^T (= A^{-k}) = Q^{(k)} (R^{(k)})^{-T}.$$
   What does this imply about the *final* column of $Q^{(k)}$? Connect it with the power method applied to a certain initial vector.

   (c) Still assuming A is symmetric, use (b) to show that in the shifted QR iteration, $A_k - s_k I = Q_k R_k, A_{k+1} = R_k Q_k + s_k I$, if the shift s_1 is chosen to be equal to an eigenvalue λ_*, then the final row of A_2 will be equal to $[0, \ldots, 0, \lambda_*]$ with probability 1.

   (d) (optional) Discuss (b),(c) in the nonsymmetric case.

4. (a) Explain how to reduce a square matrix $A \in \mathbb{R}^{n\times n}$ to an upper Hessenberg form via a sequence of orthogonal similarity transformations $A \leftarrow QAQ^T$ involving Householder reflectors. How does the process simplify if A was symmetric?

   (b) Why is it not possible to reduce A directly to a triangular (or diagonal) form?

   (c) Examine the complexity of one step of the QR algorithm for the reduced matrix (both nonsymmetric and symmetric cases). Is it $O(n^3), O(n^2)$ or $O(n)$?

5. Let $A, B \in \mathbb{R}^{n\times n}$ be symmetric. Consider the generalised eigenvalue problem $Ax = \lambda Bx$, $x(\neq 0) \in \mathbb{C}^n, \lambda \in \mathbb{C}$.
   (a) Show by an example that it is possible for a (generalised) eigenvalue λ to be nonreal.
   (b) Suppose further that B is positive definite ($B \succ 0$). Show that all eigenvalues are real.
   (c) Continuing to assume $B \succ 0$, suggest an algorithm for solving the problem using the Cholesky factorisation. (hint: $\det(A - tB) = 0$ if and only if $\det(X(A - tB)Y) = 0$ for nonsingular $X, Y \in \mathbb{R}^{n\times n}$)

6. "eig(AB)=eig(BA)"
   (a) First, let A, B be $n \times n$ square matrices, where $\det(B) \neq 0$. Prove that the eigenvalues of AB and BA are the same.
   (b) Next let $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times m}$ ($m \geq n$). Prove that $\begin{bmatrix} AB & A \\ 0 & 0_{n\times n} \end{bmatrix}$ and $\begin{bmatrix} 0_{m\times m} & A \\ 0 & BA \end{bmatrix}$ are similar. Hence show that the nonzero eigenvalues of AB and BA are again the same.
   (c) Do you see a connection between this problem and the QR algorithm?

7. (Computational) Implement the QR algorithm in MATLAB or Python. (ideally with shifts, but this exercise can be done without introducing shifts at every iteration. Should be doable in about 10 lines of code!)
   (a) Verify 3(c,d) in your code with a 4×4 random example.
   (b) What will happen if the QR algorithm (without shifts) is applied to an orthogonal matrix A? Explain why this is not a 'counterexample' for the convergence of the QR algorithm. Modify the algorithm so that it computes an eigenvalue decomposition of A.

8. (Optional) Explain how to use Householder reflectors from left and right to reduce $A \in \mathbb{R}^{m\times n}$ ($m \geq n$) to bidiagonal form B.
   (a) For the purpose of computing the SVD, explain why we are 'allowed' to apply different orthogonal transformations from the left and right.
   (b) Explain how the SVD of A can be obtained via the symmetric eigenvalue problem with respect to B^TB. (This is not the recommended algorithm in practice)

Yuji Nakatsukasa: nakatsukasa@maths.ox.ac.uk