Qual NLA Prep Sols

Mathew Tucker

NLA

1. **Let the eigenvalues of an $n \times n$ real symmetric matric A be ordered from the largest to the smallest. Prove that for any $1 \leq k \leq n$,**

$$\lambda_k = \max_{S^k} \min_{0 \neq x \in S^k} r(x, A),$$

where S^k is any k dimensional subspace of $\mathbb{R}^n$, and $r(x, A)$ is the Rayleigh quotient $\frac{x^T A x}{x^T x}$.

Reread Wiki's Proof of Min-max Theorem

2.

(a) **Prove that the growth factor $\rho = \frac{\|U\|_{max}}{\|A\|_{max}}$ is unbounded for LU factorization without pivoting.**

Let

$$A = \begin{pmatrix} \epsilon & 1 \\ 1 & 0 \end{pmatrix}.$$

Since

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{c}{a} & 1 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d - \frac{bc}{a} \end{pmatrix}$$

we have that

$$A = \begin{pmatrix} 1 & 0 \\ \frac{1}{\epsilon} & 1 \end{pmatrix} \begin{pmatrix} \epsilon & 1 \\ 0 & -\frac{1}{\epsilon} \end{pmatrix}.$$

Hence, defining $\|M\|_{max} = \max_{i,j} |m_{i,j}|$, then as $\rho = \frac{\|u\|_{max}}{\|A\|_{max}}$ we get that

$$\rho = \frac{\frac{1}{\epsilon}}{1} = \frac{1}{\epsilon}.$$

But, ϵ is arbitrarily small, and so as $\epsilon \to \infty$ we find $\rho \to \infty$ and is therefor unbounded.

(b) **Prove that the growth factor is bounded by 2^{m-1} for LU factorization of $A \in \mathbb{R}^{m\times m}$ with row pivoting.**

We have

$$\underbrace{L_{m-1} \cdots L_2 L_1}_{L^{-1}} A = U \implies A = LU$$

For all $A \in \mathbb{R}^{m\times m}$, the $\max\left(\frac{\|L\|\|U\|}{\|A\|}\right)$ is unbounded without pivoting where $\left|(L_k)_{ik}\right| \le 1$. This implies that

$$\|A_k\| \le 2\|A_{k-1}\|_{max} \implies \|A_{m-1}\|_{max} \le 2^{m-1}\|A\|_{max}.$$

Therefore, the growth factor $\rho = \frac{\|L\|_{max}\|U\|_{max}}{\|A\|_{max}} \le 2^{m-1}$.

3. **A is a diagonalizable matrix with one eigenvalue being -1, and others residing in the unit disk centered at 2 in the complex plane. Prove that the solution to $Ax = b$ through GMRES algorithm has error**

$$\|e_n\| \le K2^{-n}$$

for some constant K.

Theorem 35.2 states,

> at step n of the GMRES iteration, the residual r_n satisfies,
>
> $$\frac{\|r_n\|}{\|b\|} \le \inf_{p_n \in P_n} \|p_n(A)\| \le \kappa(V) \inf_{p_n \in P_n} \|p_n\|_{\Lambda(A)} \le \kappa(v) \inf_{p \in P_n} \max_{\lambda \in \Lambda(A)} |p(\lambda)|.$$
>
> where $\Lambda(A)$ is the set of eigenvalues of A, V is a nonsingular matrix of eigenvectors, and $\|p_n\|_{\Lambda(A)}$ is defined by $\|p\|_S = \sup_{x \in S} |p(z)|$.

We can choose a polynomial

$$\hat{p}(z) = (1+z)q(z) = (1+z)\left(1 - \frac{z}{2}\right)^{n-1}$$

so that

$$\max_{1 \le \lambda \le 3} |q(\lambda)| \le \left(1 - \frac{1}{2}\right)^{n-1}$$

or equivalently

$$\max_{|\lambda - 2| \le 1} |q(\lambda)| \le 2^{1-n}.$$

Thus we have

$$\|r_n\| \le \kappa(V)\|b\| \max_{\lambda \in \Lambda(A)} |p(\lambda)| = \kappa(V)\|b\| \max_{|\lambda-2| \le 1} |\lambda+1||q(\lambda)| \le \kappa(V)\|b\| \max_{|\lambda-2| \le 1} |\lambda+1||q(\lambda)|$$

$$\le \kappa(V)\|b\||3+1|2^{1-n} = \kappa(V)\|b\|2^{3-n}$$

Finally,

$$\|e_n\| = \left\|A^{-1}r_n\right\| \le \left(8\ \kappa(V)\|b\|\left\|A^{-1}\right\|\right)2^n,$$

and as $\left(8\ \kappa(V)\|b\|\left\|A^{-1}\right\|\right)$ is constant we have our proof.

4.

(a) **State and prove the Bauer-Fike Theorem.**

The Bauer-Fike Theorem states:

$$\text{if } A = XDX^{-1}, \text{ then } \text{dist}(\lambda(A+B),\ \Lambda(A)) \leq \kappa(X)\|B\|.$$

To prove, let's assume $(A+B)x = \lambda x$. Let $x = Xy$ and $C = X^{-1}BX$, then we find

$$(A+B)x = \lambda x$$

$$= \left(XDX^{-1}+B\right)x = \lambda x$$

$$= \left(XDX^{-1}+B\right)Xy = \lambda Xy$$

$$= X^{-1}\left(XDX^{-1}+B\right)Xy = \lambda y$$

$$= \left(DX^{-1}+X^{-1}B\right)Xy = \lambda y$$

$$= \left(D+X^{-1}BX\right)y = \lambda y$$

$$= (D+C)y = \lambda y$$

$$\Longrightarrow Cy = \lambda y - Dy = (\lambda I - D)y$$

Hence we have that,

$$\text{dist}(\lambda, \Lambda(A))\|y\| \leq \|(\lambda I - D)y\| = \|Cy\| \leq \|C\|\|y\|$$

yielding

$$\text{dist}(\lambda, \Lambda(A)) \leq \|C\|.$$

And as $C = X^{-1}BX$ we get that

$$\|C\| \leq \left\|X^{-1}\right\|\|B\|\|X\| = \kappa(X)\|B\|.$$

Hence,

$$\text{dist}(\lambda, \Lambda(A))\|y\| \leq \kappa(X)\|B\|$$

and our proof. ■

(b) **Show that the eigenvalue problem for Hermitian matrices is well-conditioned.**

On pg 199????

(c) **Give an example that this is not true for the non-Hermitian matricies.**

$$\begin{pmatrix} 4 & 1 \\ 0 & 1 \end{pmatrix}$$

5. **Let**

$$P = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}, \qquad A = P\begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}.$$

Determine the least square solution to the over-determined linear system $Ax = b$.

First lets find A,

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 3 \\ 1 & -1 \end{pmatrix}.$$

Now, we need to find the solution to the normal equation, $A^*Ax = A^*b$. Since A is full rank, the solution to the least squares problem is $x = \left(A^*A\right)^{-1}A^*b$.

Let's solve this in parts:
First,

$$A^*A = \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 3 \\ 1 & -1 \end{pmatrix}^T \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 3 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ 4 & 20 \end{pmatrix}$$

so,

$$\left(A^*A\right)^{-1} = \frac{1}{\det(A^*A)} \operatorname{adj}\left(A^*A\right) = \frac{1}{64}\begin{pmatrix} 20 & -4 \\ -4 & 4 \end{pmatrix}.$$

Next,

$$A^*b = \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 3 \\ 1 & -1 \end{pmatrix}^T \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \end{pmatrix}.$$

Hence we have that,

$$x = \frac{1}{64}\begin{pmatrix} 20 & -4 \\ -4 & 4 \end{pmatrix}\begin{pmatrix} 6 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{7}{4} \\ \\ \frac{-1}{4} \end{pmatrix}.$$

6. **Prove that the convergence of Rayleigh quotient iteration for a hermitian matrix is ultimately cubic.**

Recall, normal matrix $A \in \mathbb{R}^{m\times m}$ is orthogonally diagonalizable and the Rayleigh quotient of a vector $x \in \mathbb{R}^m$ is the scalar

$$r(x,A) = r(x) = \frac{x^T A x}{x^T x}.$$

Observe, $r(x,A) = \lambda$, λ an eigenvalue of A, if x is an eigenvector of A.

Let q_i be orthonormal eigenvectors of A with distinct eigenvalues. Then any vector x_n can be written as $x_n = \sum_{i=1}^k a_i q_i$, the linear combination of q_i. As $Aq_l = \lambda_l q_l$ we get that,

$$r(x_n) = \frac{a_1^2\lambda_1 + \cdots + a_k^2\lambda_k}{a_1^2 + \cdots + a_k^2} = \frac{\sum_{i=1}^k a_i^2 \lambda_i}{\sum_{i=1}^k a_i^2}.$$

Assuming x_n converges to q_1, then $a_1 \approx 1$ and, for $i \neq 1$, $a_i = O(\theta)$ where θ is small; and, we get that,

$$r(x_n) = \frac{a_1^2\lambda_1 + \cdots + a_k^2\lambda_k}{a_1^2 + \cdots + a_k^2} = \frac{a_1^2\lambda_1 + O(\theta^2)}{a_1^2 + O(\theta^2)} = \lambda_1 + O(\theta^2).$$

Applying the Rayleigh quotient iteration we get,

$$x_{n+1} = (A - r(x_n)I)^{-1} x_n = \sum_{i=1}^k \frac{a_i}{\lambda_i - r(x_n)} q_i = \frac{a_1}{O(\theta^2)} q_1 + \cdots + \frac{O(\theta)}{\lambda_k - \lambda_1 + O(\theta^2)} q_k$$

which is parallel to

$$a_1 q_1 + \cdots + \frac{O(\theta^3)}{\lambda_k - \lambda_1} q_k.$$

And after nomalization, $\|x_{n+1} - x_n\| \to \|x_{n+1} - q_1\| = O(\theta^3)$, hence we have cubic convergence.

7. **Suppose A is a real symmetric matrix with eigenvalues more or less uniformly distributed over $[2,18]$ together with an outlier at $\lambda = 50$. How many steps of the conjugate gradient iteration must be taken to be sure of reducing the initial error $\|e_0\|_A$ by a factor of 2^{20}?**

   By theorem 38.5, if all the eigenvalues are in $[2,18]$, then the A-norms of the error satisfy,

$$\frac{\|e_n\|_A}{\|e_0\|_A} \le 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n.$$

   Here $\kappa = \frac{18}{2} = 9$ and so,

$$\frac{\|e_n\|_A}{\|e_0\|_A} \le 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n = 2\left(\frac{\sqrt{9}-1}{\sqrt{9}+1}\right)^n = 2\left(\frac{1}{2}\right)^n = \left(\frac{1}{2}\right)^{n-1}.$$

   Including $\lambda = 50$ we get that $\frac{\|e_n\|_A}{\|e_0\|_A} \le (2)2^{1-n}$. And finally, setting $2^{-20} = 2^{2-n}$ we find $n = 22$

8. **Derive the asymptotic operation count of Gaussian elimination applied on an $m \times m$ real matrix A.**

Per Trefethen(pg 151), as the work of the algorithm is dominated by the inner most loop, $u_{j,k:m} = u_{j,k:m} - l_{j,k} u_{k,k:m}$, it is sufficient to consider just this part. Observe, for the k-iteration of our algorithm, we are only acting on $m - k + 1$ elements (the "plus one" is an artifact of counting) twice. Hence we will need to take a double sum, representing each loop, of $2(m - k + 1)$. As our outer loop is for $k = 1$ to $m - 1$ and our inner loop is for $j = k + 1$ to m we have that the number of flops is

$$\sum_{k=1}^{m-1} \sum_{j=k+1}^{m} 2(m-k+1).$$

Reindexing our inner summation with $j = 1$ we find

$$\sum_{k=1}^{m-1} \sum_{j=1}^{m-k} 2(m-k+1) \sim \sum_{k=1}^{m-1} \sum_{j=1}^{m-k} 2(m-k) = \sum_{k=1}^{m-1} 2(m-k)(m-k) \sim \sum_{k=1}^{m-1} 2k^2 = 2\frac{(m-1)m(2m+1)}{6} \sim \frac{2m^3}{3} \text{ flops.}$$

9. **Given**

$$A = \begin{pmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{pmatrix}, \qquad \epsilon = 10^{-9}$$

(a) **Find A^*A and the 2-norm $\kappa(A)$.**

As ϵ is real,

$$A^* = A^T = \begin{pmatrix} 1 & \epsilon & 0 & 0 \\ 1 & 0 & \epsilon & 0 \\ 1 & 0 & 0 & \epsilon \end{pmatrix}$$

and we get that

$$A^*A = \begin{pmatrix} 1 & \epsilon & 0 & 0 \\ 1 & 0 & \epsilon & 0 \\ 1 & 0 & 0 & \epsilon \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{pmatrix} = \begin{pmatrix} 1+\epsilon^2 & 1 & 1 \\ 1 & 1+\epsilon^2 & 1 \\ 1 & 1 & 1+\epsilon^2 \end{pmatrix}$$

which can be decomposed as the sum of two matrix,

$$\begin{pmatrix} 1+\epsilon^2 & 1 & 1 \\ 1 & 1+\epsilon^2 & 1 \\ 1 & 1 & 1+\epsilon^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} + \begin{pmatrix} \epsilon^2 & 0 & 0 \\ 0 & \epsilon^2 & 0 \\ 0 & 0 & \epsilon^2 \end{pmatrix}.$$

Hence, to find the eigenvalues of A^*A it is sufficient to add ϵ to the eigenvalues of

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Solving $|A - \lambda I| = 0$ we find the eigenvalues of M are $\{3, 0, 0\}$; and, thus the eigenvalues of A^*A are $\{3+\epsilon^2, \epsilon^2, \epsilon^2\}$. Therefore, $\kappa(A) = \sqrt{\frac{3+\epsilon^2}{\epsilon^2}} = 1 + \sqrt{3}\ 10^9$.

(b) **MATLAB returns** $\text{rank}(A) = 3$**, but** $\text{rank}(A^*A) = 1$**. Explain.**

The reason why is because ϵ^2 is on the order of machine error, so the computer reports it as 0. As a result, the matrix $M + \epsilon^2 I$ is recorded as M which is a matrix of rank 1.

10. **Derive the asymptotic operation count for the following algorithms applied on a full-rank $m \times n\,(n \le m)$ matrix A.**

(a) **Reduced QR factorization by modified Gram-Schmidt orthogonalization.**

Per Trefethen(pg 59), as the work of the algorithm is dominated by the inner most loop,

$$r_{i,j} = q^* v_j$$

$$v_j = v_j - r_{i,j} q_i,$$

it is sufficient to consider just this part. Observe, the first line is an inner product which requires m multiplications and $m-1$ additions. The second line, $v_j = v_j - r_{i,j} q_i$, requires m multiplications and m subtractions. Hence the total flops for a single iteration of the inner most loop is $\sim 4m$. We will now need to take a double sum, representing each loop, of $4m$. As our outer loop is for $i = 1$ to n and our inner loop is for $j = i+1$ to n we have that the number of flops is

$$\sum_{i=1}^{n} \sum_{j=i+1}^{n} 4m = \sum_{i=1}^{n} (n-i-1)4m = 4m\left(\sum_{i=1}^{n} n - \sum_{i=1}^{n} i - \sum_{i=1}^{n} 1\right) = 4m\left(n^2 - \frac{n(n+1)}{2} - n\right) = 4m\left(\frac{n^2}{2} - \frac{3n}{2}\right) \sim 2mn^2 \text{ flops.}$$

(b) **Reduced QR factorization by Householder triangularization(without forming Q).**

Per Trefethen(pg 74), as the work of the algorithm is dominated by the inner most loop, $A_{k:m,k:n} = A_{k:m,k:n} - 2v_k(v_k^* A_{k:m,k:n})$, it is sufficient to consider just this part. Observe, per iteration we have

- $(m-k)(n-k)$ flops for the scalar product
- $2(m-k)(n-k)$ flops for the dot product
- $(m-k)(n-k)$ flops for the subtraction,

hence the total flops for a single iteration of the inner most loop is $4(m-k)(n-k)$. We now need only take a single sum, representing the outer most loop. As our outer loop is from $k=1$ to n we have that the number of flops is

$$\sum_{k=1}^{n} 4(m-k)(n-k) = 4\left(\sum_{k=1}^{n} mn - \sum_{k=1}^{n} k(m+n) + \sum_{k=1}^{n} k^2\right) = 4\left(mn^2 - \frac{n(n+1)(m+n)}{2} + \frac{n(n+1)(2n+1)}{6}\right)$$

$$= 4mn^2 - 2n(n+1)(m+n) + \frac{2n(n+1)(2n+1)}{3} = 2mn^2 - 2mn - \frac{2n^3+2n}{3} \sim 2mn^2 - \frac{2n^3}{3} \text{ flops.}$$