

Homework 5: Report

Dante Buhl

Feb. 26th 2024

2 Numerical Problems

1.

3 Theory Problems

1. Consider the Householder matrix defined by

$$H = I - 2 \frac{vv^T}{v^T v}$$

- (a) Show that for any nonzero vector v , the matrix is orthogonal and symmetric.
- (b) Let a be any nonzero vector and let $v = a + \alpha e_1$, where $\alpha = \text{sign}(a_{11}) \|a\|_2$. Show that $Ha = -\alpha e_1$ by direct calculation.
- (c) Determine v and α that transforms,

$$H \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

- (d) Given the vector $a = (2, 3, 4)^T$, specify a Householder transformation that annihilates the third component of a .
- (e) What are the eigenvalues of H for any nonzero vector x ?
- (a) *Proof.* Begin by looking at the transpose of H .

$$H^T = \left(I - 2 \frac{vv^T}{v^T v} \right)^T$$

$$H^T = I - \frac{2}{v^T v} (vv^T)^T = I - 2 \frac{vv^T}{v^T v} = H$$

So we can immediately see that H is symmetric. Next we look at $H^T H$.

$$H^T H = H^2 = \left(I - 2 \frac{vv^T}{v^T v} \right) \left(I - 2 \frac{vv^T}{v^T v} \right)$$

$$H^T H = I - 2 \frac{vv^T}{v^T v} - 2 \frac{vv^T}{v^T v} + \frac{4}{(v^T v)^2} vv^T vv^T$$

$$H^T H = I - 4 \frac{vv^T}{v^T v} + \frac{4v^T v}{(v^T v)^2} vv^T = I$$

$$H^T H = I$$

So we have that H is orthogonal as well. □

(b) *Proof.*

$$v = a + \alpha e_1, \quad v^T = a^T + \alpha e_1^T$$

$$H = I - 2 \frac{vv^T}{v^T v}$$

$$= I - \frac{2}{a^T a + 2\alpha a_1 + \alpha^2} \left(aa^T + \begin{bmatrix} \frac{\alpha a^T}{0} \\ \vdots \\ 0 \end{bmatrix} + [\alpha a \mid 0 \mid \cdots \mid 0] + \begin{bmatrix} \alpha^2 & 0 & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \right)$$

We now multiply by a .

$$Ha = a - \frac{2}{a^T a + 2\alpha a_1 + \alpha^2} \left(aa^T a + \begin{bmatrix} \frac{\alpha a^T}{0} \\ \vdots \\ 0 \end{bmatrix} a + [\alpha a \mid 0 \mid \cdots \mid 0] a + \begin{bmatrix} \alpha^2 & 0 & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} a \right)$$

Recall that $a^T a = \alpha^2$.

$$Ha = a - \frac{2}{\alpha^2 + 2\alpha a_1 + \alpha^2} (\alpha^2 a + \alpha^3 e_1 + \alpha a_1 a + \alpha^2 a_1 e_1)$$

$$Ha = a - \frac{\alpha}{\alpha(\alpha + a_1)} ((\alpha + a_1)a + \alpha(\alpha + a_1)e_1)$$

$$Ha = a - a - \alpha e_1 = -\alpha e_1$$

$$Ha = -\alpha e_1$$

□

(c) *Proof.* We have evidently that $a = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$. Therefore we have that $\alpha = -\sqrt{4}$ and $v = a + \sqrt{4}e_1$.

(Note that in this specific case, I am not using α as the exact modifier in the v vector but rather $-\alpha$.)

$$v = \begin{bmatrix} 1 + \sqrt{4} \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

□

(d) *Proof.* In order to eliminate the third component of a we simply chose the vector $v = a + \|a\|_2 e_3$ for the Householder transformation and then the following changes. We will have that,

$$\left(I + \frac{4}{\|a\|_2^2} H \right) a = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$$

□

(e)

2. The Schur decomposition theorem states that every square matrix $A \in \mathbb{C}^{m \times m}$ has a Schur Decomposition, $A = QUQ^*$, where Q is a unitary and U is upper triangular. Use this theorem to prove that, for an arbitrary norm $\|\cdot\|$,

$$\lim_{n \rightarrow \infty} \|A^n\| = 0 \iff \rho(A) < 1$$

(Note: Show the claim first with the 2-norm or the Frobenius norm and use the fact that all norms are equivalent in a finite vector space.)

Proof. (\implies)

Let us look at the two-norm of A^n . Next we investigate an arbitrary eigenvalue, eigenvector pair of A , (λ, v) . We have that,

$$\|A^n\|_2 \geq \frac{\|A^n v\|_2}{\|v\|_2} = \frac{|\lambda|^n \|v\|_2}{\|v\|_2} = |\lambda|^n$$

We go back to the limit as $n \rightarrow \infty$.

$$\lim_{n \rightarrow \infty} \|A^n\|_2 = 0 > |\lambda|^n$$

Therefore we have that $|\lambda| < 1$. Since this observation is independent of which p-norm is chosen and which eigenvalue of A we chose, we must have that all eigenvalues of A are such that $|\lambda| < 1$ and therefore we must have that $\rho(A) < 1$.

(\impliedby)

Given that for a matrix $A = QUQ^*$ and $\rho(A) < 1$. We have that the eigenvalues of A are along the diagonal of U (Mini proof: A is similar to U , and $\det(U - \lambda I) = \prod_{i=1}^m (u_{ii} - \lambda)$ since U is upper triangular. Therefore the eigenvalues of U are found along its diagonal.) Next since $\rho(A) < 1$ we have that all of the eigenvalues of A are of absolute value less than 1. So we have that,

$$\lim_{n \rightarrow \infty} \|A^n\| = \lim_{n \rightarrow \infty} \|QU^nQ^*\| = \lim_{n \rightarrow \infty} \|U^n\|$$

Now to show that the norm is equal to zero irrespective of which norm is taken, we need to show that $U^n \rightarrow 0$. This requires induction. We start with the base case: that the diagonal goes to zero. We have that the diagonal is constructed of the eigenvalues of A which are all less than 1 in absolute value. We look at the matrix product U^2 .

$$U^2 = UU = \begin{bmatrix} \lambda_1^2 & u_{12}(\lambda_1 + \lambda_2) & \cdots \\ 0 & \lambda_2^2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Notice that the diagonal entries of U^n will be $u_{ii} = \lambda_i^n$. Since all $|\lambda_i| < 1$ we have that $\lambda^n \rightarrow 0$ as $n \rightarrow \infty$ or rather $\lambda^n \approx 0$ as n becomes large but still finite. So we have the the diagonal will tend towards zero as n increases. Our inductive step is then to look at U described as the following.

$U_k = \begin{bmatrix} \mathbf{0} & R_k \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$. Here R_k is an upper triangular matrix, $R_k \in \mathbb{R}^{m-k \times m-k}$. Our goal is to show that by

multiplying by U some discrete amount of times, that we will find U_{k+1} where, $U^{k+1} = \begin{bmatrix} \mathbf{0} & R_{k+1} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$ and $R_{k+1} \in \mathbb{R}^{m-k-1 \times m-k-1}$. That is, the diagonal of R_k will go to zero, producing R_{k+1} inside of it. We begin,

$$U_k U = \begin{bmatrix} \mathbf{0} & r_{11}\lambda_k & \cdots & \cdots \\ \mathbf{0} & 0 & r_{22}\lambda_{k+1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

We can see that if we multiply by U by a finite amount of times, we will eventually obtain similarly that $U_k U^n$ which has zeros on what used to be the diagonal of R_k (the diagonal of R_k will be multiplied by its respective λ some amount of times such that, $r_{jj}\lambda_{k+j}^n \rightarrow 0$ as $n \rightarrow \infty$. Notice that for some smaller, finite n we will have $r_{jj}\lambda_{k+j}^n \approx 0$. Therefore after multiplying by some finite number of

U we will effectively obtain U_{k+1} which has the following form $U_{k+1} = \begin{bmatrix} \mathbf{0} & R_{k+1} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$ and $R_{k+1} \in \mathbb{R}^{m-k-1 \times m-k-1}$. Notice that this is effectively done with finite multiples of U . That is, $U_{k+1} = U^n$ such that $n < \infty$.

The inductive step is complete! If we have U_k from finite multiples of U , through additional finite multiples of U we obtain U_{k+1} . Therefore, we return to the limit and obtain the matrix U_m which with all intents and purposes is an $m \times m$ zero matrix to machine precision.

$$\lim_{n \rightarrow \infty} \|A^n\| = \lim_{n \rightarrow \infty} \|U^n\| = \|U_m\|$$

Finally we have that a zero matrix has norm of zero.

$$\lim_{n \rightarrow \infty} \|A^n\| = \|U_m\| = \|\mathbf{0}\| = 0$$

□

3. Let $A \in \mathbb{C}/\mathbb{R}^{m \times n}$ and $B \in \mathbb{C}/\mathbb{R}^{n \times m}$. Show that the matrices $\begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}$ have the same eigenvalues.

Proof. We start by looking at the determinant of block matrices. Take for example the matrix $\Gamma = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$.

$$\Gamma = \begin{bmatrix} I & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} I & B \\ 0 & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}$$

We look at the determinant of Γ .

$$\det(\Gamma) = \det\left(\begin{bmatrix} I & 0 \\ 0 & D \end{bmatrix}\right) \det\left(\begin{bmatrix} I & B \\ 0 & I \end{bmatrix}\right) \det\left(\begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}\right)$$

$$\det(\Gamma) = \det(D) \det(A)$$

The same is obviously true for a matrix Γ of the form, $\Gamma = \begin{bmatrix} A & 0 \\ B & D \end{bmatrix}$ with A invertible ($\Gamma = \begin{bmatrix} I & 0 \\ BA^{-1} & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}$). We now look at the eigenvalues of $M_1 = \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix}$ and $M_2 = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}$. We look at the determinants,

$$\det(M_1 - \lambda I) = \det(AB - \lambda I) \det(-\lambda I)$$

$$\det(M_2 - \lambda I) = \det(-\lambda I) \det(BA - \lambda I)$$

Notice that if λ is an eigenvalue of its respective matrix then this determinant product must be equal to zero. We then take the case of each determinant can be zero. For M_1 we have that its eigenvalues are either zero, or the eigenvalues of AB by definition.

$$\det(M_1 - \lambda I) = 0, \quad \det(AB - \lambda I) = 0, \quad \det(-\lambda I) = -\lambda = 0$$

Similarly we have that the eigenvalues for M_2 are either zero or the eigenvalues of BA . Notice from a different homework problem (hw2), we have that the matrix products AB and BA have the same eigenvalues. This is because,

$$ABv = \lambda v, \quad Bv = y, \implies BAy = \lambda y$$

Finally both M_1 and M_2 have eigenvalues of zero and the eigenvalues of AB/BA .

□

4. Show that for a real-valued square matrix the Gerschgorin theorem also holds with the bounds r_i which are given by the partial column sums (instead of the partial row sums):

$$r_i = \sum_{j=1, j \neq i}^m |a_{i,j}|, \quad i = 1, \dots, m$$

Proof. We begin by reviewing Gershgorin's Theorem. We have for a square matrix A . It's eigenvalues are bound by the row sums along A . That is,

$$|\lambda - a_{ii}| \leq \sum_{j=1, j \neq i}^m |a_{ij}|$$

We now look at some properties of A . We have that since $\det(A) = \det(A^T)$ for real matrices A that, $\det(A - \lambda I) = \det(A^T - \lambda I^T) = \det(A^T - \lambda I)$. We notice that A and A^T have the same eigenvalues. We next apply Gershgorin's theorem for A^T . We find,

$$|\lambda - a_{ii}| \leq \sum_{j=1, j \neq i}^m |a_{ji}|$$

Notice that we have,

$$\begin{aligned} |\lambda - a_{ii}| &\leq \sum_{j=1, j \neq i}^m |a_{ij}|, \quad |\lambda - a_{ii}| \leq \sum_{j=1, j \neq i}^m |a_{ji}| \\ |\lambda - a_{ii}| &\leq \min \left\{ \sum_{j=1, j \neq i}^m |a_{ij}|, \sum_{j=1, j \neq i}^m |a_{ji}| \right\} \end{aligned}$$

That is, Gershgorin's Theorem applies for the column sums as well in addition to the row sums and often you can minimize the radius by picking the lower of the column or row sum. \square

5. Use the Gerschgorin theorem to show that the following matrix has exactly one eigenvalue in each of the four circles: $|z - k| \leq 0.1$, $k = 1, 2, 3, 4$.

$$A = \begin{bmatrix} 1.0 & 0.3 & 0.1 & 0.4 \\ 0.0 & 2.0 & 0.0 & 0.1 \\ 0.0 & 0.4 & 3.0 & 0.0 \\ 0.1 & 0.0 & 0.0 & 4.0 \end{bmatrix}$$

Proof. We simply apply Gershgorin's Theorem.

$$\begin{aligned} |\lambda_1 - 1.0| &\leq \min \left\{ \sum_{j=1, j \neq 1}^m |a_{1j}|, \sum_{j=1, j \neq 1}^m |a_{j1}| \right\} \\ |\lambda_2 - 2.0| &\leq \min \left\{ \sum_{j=1, j \neq 2}^m |a_{2j}|, \sum_{j=1, j \neq 2}^m |a_{j2}| \right\} \\ |\lambda_3 - 3.0| &\leq \min \left\{ \sum_{j=1, j \neq 3}^m |a_{3j}|, \sum_{j=1, j \neq 3}^m |a_{j3}| \right\} \\ |\lambda_4 - 4.0| &\leq \min \left\{ \sum_{j=1, j \neq 4}^m |a_{4j}|, \sum_{j=1, j \neq 4}^m |a_{j4}| \right\} \end{aligned}$$

$$|\lambda_1 - 1.0| \leq 0.1$$

$$|\lambda_2 - 2.0| \leq 0.1$$

$$|\lambda_3 - 3.0| \leq 0.1$$

$$|\lambda_4 - 4.0| \leq 0.1$$

Since the circles don't overlap, we have that A has for non-zero eigenvalues each within the circles described above. \square

6. Let $A \in \mathbb{R}^{m \times m}$ be real and symmetric that is positive definite. Let $y \in \mathbb{R}^m$ be nonzero. Prove that the limit exists and is an eigenvalue of A .

$$\lim_{k \rightarrow \infty} \frac{y^T A^{k+1} y}{y^T A^k y}$$

Proof. First, since A is real and symmetric, we can decompose A into an eigenvector matrix, full with orthonormal eigenvectors which form a basis for \mathbb{R}^m and a diagonal matrix with eigenvalues of A on the diagonal. Since its columns form a basis for \mathbb{R}^m we have that we can express y as a linear combination of the eigenvectors of v : $y = c_1 v_1 + \dots + c_m v_m$. We now define λ_* such that, $|\lambda_*| \geq |\lambda_i|, \forall \lambda_i \in \Lambda(A), c_* \neq 0$.

$$A = V D V^{-1}, \quad y = c_* v_* + c_1 v_1 + \dots + c_m v_m$$

$$A^n = V D^n V^{-1}, \quad A^n y = c_* \lambda_*^n v_* + c_1 \lambda_1^n v_1 + \dots + c_m \lambda_m^n v_m$$

$$A^n y = \lambda_*^n \left(c_* v_* + c_1 \frac{\lambda_1^n}{\lambda_*^n} v_1 + \dots + c_m \frac{\lambda_m^n}{\lambda_*^n} v_m \right)$$

Note from our definition of λ_* we either have that $|\lambda_*| \geq |\lambda_i| \vee (|\lambda_*| < |\lambda_i| \wedge c_i = 0)$. Thereby, we have that

$$\lim_{n \rightarrow \infty} A^n y = \lambda_*^n c_* v_*$$

Therefore we have,

$$\lim_{k \rightarrow \infty} \frac{y^T A^{k+1} y}{y^T A^k y} = \frac{\lambda_*^{k+1} c_* y^T v_*}{\lambda_*^k c_* y^T v_*} = \lambda_*$$

Where λ_* is the largest eigenvalue of A which has an eigenvector as part of the expansion of y into the basis of the eigenvectors of A . \square

7. Let $A \in \mathbb{R}^{m \times m}$ be real with nonnegative entries such that

$$\sum_{j=1}^m a_{ij} = 1 \quad (1 \leq i \leq m)$$

Prove that no eigenvalue of A has an absolute value greater than 1.

Proof. We start by looking at an arbitrary p-norm of the matrix product of A and one of its eigenvectors.

$$Av = \lambda v, \quad \|Av\| = \|\lambda v\|$$

$$\|Av\| \leq \|A\| \|v\|, \quad |\lambda| \|v\| \leq \|A\| \|v\|$$

$$|\lambda| \leq \|A\|$$

We obtain the following inequality which is satisfied by any p-norm we would like to take. We simply chose $p = 1$, the one norm. By definition of the one norm for matrices (lecture 5 notes) we have,

$$|\lambda| \leq \|A\|_1 = \max_i \sum_{j=1}^m a_{ij}$$

We also have that A is given such that its row sum along any row will sum to 1. Therefore we have,

$$|\lambda| \leq \|A\|_1 = 1, \implies |\lambda| \leq 1$$

\square

8. Let $A \in \mathbb{R}^{m \times m}$ be a non-defective matrix with its eigenvalues $\{\lambda_i\}_{i=1}^m$ and its singular values $\{\sigma_i\}_{i=1}^m$, satisfying

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_m|$$

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m$$

Let $\rho(A)$ be the spectral radius of A and $\text{cond}(A) = \|A\|_2 \|A^{-1}\|_2$ be the condition number of A . Let A be normal, i.e., $A^T A = A A^T$. Show that:

- (a) $\sigma_i = |\lambda_i|, 1 \leq i \leq m$.
(b) $\|A\|_2 = |\lambda_1| = \rho(A)$.

- (a) *Proof.* We begin with the fact that A is non-defective, is normal, and then has a Singular Value Decomposition. First we look at a diagonalization of A by fact that A is non-defective.

$$V^{-1} A V = D_\lambda \implies A = V_\lambda D_\lambda V_\lambda^{-1}$$

Notice that since, A is non-defective, the eigenvalues of A form a basis for \mathbb{R}^m (i.e. V_λ is orthogonal). More importantly, as is given in the diagonalization V_λ is invertible.

$$V^T V = I, \implies V^T V V^{-1} = V^{-1}, \implies V^T = V^{-1}$$

$$A^T A = V^{-T} D V^T V D V^{-1} = V D^2 V^{-1} = A^2$$

We also have that A is invertible since it is nondefective. So we find that A is symmetric.

$$A^T A A^{-1} = A^2 A^{-1} \implies A^T = A$$

At this point we can reference Theorem 5.5 of the textbook (on page 45 or 44 of the pdf). The rest of the proof follows. The logic cited there states is to take $A = V D V^{-1}$ where $V^T = V^{-1}$ and then to create a matrix S_λ which is a diagonal matrix with only the signs of each eigenvalue on D in the diagonal. That is we can decompose $D = S_\lambda |D|$. We then write

$$A = V S_\lambda |D| V^{-1}$$

Notice that since V is orthogonal, so is $V S_\lambda$, and alternatively V^T was already orthogonal. Then notice that $|D|$ is a positive definite diagonal matrix. Hence we have that $U = V S_\lambda$, $\Sigma = |D|$, and $V^T = V^{-1}$ is a valid SVD of A . Therefore we have that the singular values of A are $\sigma_i = |\lambda_i|$ with the implicit scaling (ordering of λ_i) $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_m|$.

□

- (b) *Proof.* If it is given that for a matrix A that is a square, normal, and non-defective matrix, we have that $\sigma_i = |\lambda_i|, \forall i, 1 \leq i \leq m$. Note that both σ_i and $|\lambda_i|$ monotonically decrease as i increases in this particular problem. Therefore we then look at the two-norm of A . We have from some lecture note (probably lecture 5) the following property of the two-norm. $\|A\|_2 = \sigma_1$. Then we also have that as is given and by definition of the spectral radius of A , $\rho(A) = |\lambda_1|$, the largest eigenvalue of A .

$$\|A\|_2 = \sigma_1 = |\lambda_1| = \rho(A)$$

□