APPM 5610 - Homework 4

Eappen Nelluvelil

February 16, 2022

1. Show that Jacobi's method for finding eigenvalues of a real symmetric matrix is ultimately quadratically convergent. Assume that all off-diagonal elements of the matrix $\mathbf{A}^k$ are $\mathcal{O}(\epsilon)$, where k enumerates Jacobi sweeps. Show that then all rotations of the next Jacobi sweep are of the form

$$1$$

Then demonstrate that this implies that, after the sweep, all off-diagonal elements of $\mathbf{A}_{k+1}$ are $\mathcal{O}\left(\epsilon^2\right)$. Assume that all eigenvalues are non-zero and distinct.

Let $\mathbf{A}$ be a given $n \times n$ real, symmetric matrix whose eigenvalues are non-zero and distinct.

We denote Jacobi rotation matrices by $\mathbf{J}$, and we see that $\mathbf{J}$ is identical to the $n \times n$ identity matrix $\mathbf{I}$ except for in four entries: $\mathbf{J}_{pp} = c$, $\mathbf{J}_{pq} = s$, $\mathbf{J}_{qp} = -c$, and $\mathbf{J}_{qq} = s$, where p and q are the row and column, respectively, of $\mathbf{A}$ such that $\mathbf{A}_{pq}$ is the largest off-diagonal entry in magnitude; $c = \cos(\theta)$; and $s = \sin(\theta)$. Here, θ is the angle by which we rotate the two-dimensional subspace spanned by columns p and q of the n-dimensional subspace spanned by the columns of $\mathbf{A}$.

From *Matrix Computations*, the values c and s are determined from the following equations:

$$c = \frac{1}{\sqrt{1+t_{\min}^2}},$$
$$s = t_{\min} c,$$

where

$$\tau = \frac{a_{qq} - a_{pp}}{2a_{pq}},$$
$$t_{\min} = \begin{cases} \frac{1}{\tau + \sqrt{1+\tau^2}} & \text{if } \tau \geq 0 \\ \frac{1}{\tau - \sqrt{1+\tau^2}} & \text{if } \tau < 0 \end{cases}.$$

To show that Jacobi's method for finding eigenvalues is ultimately quadratically convergent, we suppose that all off-diagonal elements of $\mathbf{A}^k$ are $\mathcal{O}(\epsilon)$. From the definition of τ, we see that

$$\begin{aligned} \tau &= \frac{a_{qq} - a_{pp}}{2a_{pq}} \\ &\sim \frac{a_{qq} - a_{pp}}{2\mathcal{O}(\epsilon)} \quad \text{since the off-diagonal entries of } \mathbf{A}^k \text{ are } \mathcal{O}(\epsilon) \\ &\sim \mathcal{O}\left(\frac{1}{\epsilon}\right). \end{aligned}$$

From the definition of $t_{\min}$, we see that

$$\begin{aligned}
t_{\min} &= \begin{cases} \frac{1}{\tau+\sqrt{1+\tau^2}} & \text{if } \tau \geq 0 \\ \frac{1}{\tau-\sqrt{1+\tau^2}} & \text{if } \tau < 0 \end{cases} \\
&\sim \begin{cases} \frac{1}{\mathcal{O}\left(\frac{1}{\epsilon}\right)+\sqrt{1+\mathcal{O}\left(\frac{1}{\epsilon^2}\right)}} & \text{if } \tau \geq 0 \\ \frac{1}{\mathcal{O}\left(\frac{1}{\epsilon}\right)-\sqrt{1+\mathcal{O}\left(\frac{1}{\epsilon^2}\right)}} & \text{if } \tau < 0 \end{cases} \\
&\sim \begin{cases} \frac{1}{\mathcal{O}\left(\frac{1}{\epsilon}\right)+\sqrt{\mathcal{O}\left(\frac{1}{\epsilon^2}\right)}} & \text{if } \tau \geq 0 \\ \frac{1}{\mathcal{O}\left(\frac{1}{\epsilon}\right)-\sqrt{\mathcal{O}\left(\frac{1}{\epsilon^2}\right)}} & \text{if } \tau < 0 \end{cases} \\
&\sim \begin{cases} \mathcal{O}\left(\epsilon\right) & \text{if } \tau \geq 0 \\ \frac{1}{\mathcal{O}(\epsilon)} & \text{if } \tau < 0 \end{cases}.
\end{aligned}$$

Taking the smaller of the two values, we see that $t_{\min} = \mathcal{O}\left(\epsilon\right)$. From the definition of c, we see that

$$\begin{aligned}
c &= \frac{1}{\sqrt{1+t_{\min}^2}} \\
&\sim 1 - \frac{t_{\min}^2}{2} + \ldots \\
&\sim 1 - \mathcal{O}\left(\epsilon^2\right),
\end{aligned}$$

as desired. Then, from the definition of s, we see that

$$\begin{aligned}
s &= t_{\min} c \\
&\sim \mathcal{O}\left(\epsilon\right)\left(1 - \mathcal{O}\left(\epsilon\right)\right) \\
&\sim \mathcal{O}\left(\epsilon\right) - \mathcal{O}\left(\epsilon^3\right) \\
&\sim \mathcal{O}\left(\epsilon\right),
\end{aligned}$$

as desired. Thus, $c \sim 1 - \mathcal{O}\left(\epsilon^2\right)$ and $s \sim \mathcal{O}\left(\epsilon\right)$.

To show that the Jacobi eigenvalue algorithm, when applied to $\mathbf{A}^k$, results in $\mathbf{A}^{k+1}$ having off-diagonal entries that are $\mathcal{O}\left(\epsilon^2\right)$, we note that the first iteration of the algorithm

(a) zeroes the a_{pq} and a_{qp} entries of $\mathbf{A}^k$,

(b) leaves the diagonal entries of $\mathbf{A}^k$ unaffected, and

(c) leaves the other off-diagonal entries of $\mathbf{A}^k$ unaffected, so they remain $\mathcal{O}\left(\epsilon\right)$.

In the next iteration of the algorithm, we select columns r and s of $\mathbf{A}^k$ so as to zero out a_{rs} and a_{sr}. After doing so, we see that

(a) the diagonal entries of $\mathbf{A}^k$ remain unaffected,

(b) the (p,q) and (q,p) entries from earlier are now $\mathcal{O}\left(\epsilon^2\right)$,

(c) the (r,s) and (s,r) entries are zeroed out, and

(d) rows and columns other than r and s are unaffected, so they remain either $\mathcal{O}\left(\epsilon\right)$ or $\mathcal{O}\left(\epsilon^2\right)$.

As we continue with the iterations of the sweep, the above pattern continues, and by the end of the sweep, the diagonal entries of $\mathbf{A}^{k+1}$ are $\mathcal{O}\left(\epsilon^2\right)$.

2. Show that $\mathbf{A}$ is diagonalizable if and only if there is a positive-definite self-adjoint (Hermitian) matrix $\mathbf{H}$ such that $\mathbf{H}^{-1}\mathbf{A}\mathbf{H}$ is normal.

($\Longrightarrow$) Suppose that $\mathbf{A}$ is diagonalizable, i.e., there exists an invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$. We can then use $\mathbf{P}$'s polar decomposition to write $\mathbf{P} = \mathbf{S}\mathbf{U}$, where $\mathbf{S}$ is positive-definite, Hermitian matrix and $\mathbf{U}$ is unitary. Here, the positive-definiteness of $\mathbf{S}$ is guaranteed since $\mathbf{P}$ is invertible. Thus,

$$\begin{aligned} \mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1} &\implies \mathbf{A} = \mathbf{S}\mathbf{U}\mathbf{D}\mathbf{U}^*\mathbf{S}^{-1} \\ &\implies \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{U}\mathbf{D}\mathbf{U}^*. \end{aligned}$$

Let $\mathbf{B} = \mathbf{U}\mathbf{D}\mathbf{U}^*$. We see that

$$\begin{aligned} \mathbf{B}^*\mathbf{B} &= \mathbf{U}\mathbf{D}^*\mathbf{U}^*\mathbf{U}\mathbf{D}\mathbf{U}^* \\ &= \mathbf{U}\mathbf{D}^*\mathbf{D}\mathbf{U}^* \\ &= \mathbf{U}\mathbf{D}\mathbf{U}^*\mathbf{U}\mathbf{D}^*\mathbf{U}^* \\ &= \mathbf{U}\mathbf{D}\mathbf{D}^*\mathbf{U}^* \\ &= \mathbf{B}\mathbf{B}^* \end{aligned}$$

since diagonal matrices commute. Thus, there exists a positive-definite self-adjoint matrix $\mathbf{S}$ such that $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ is normal.

($\Longleftarrow$) Suppose that there exists a positive-definite self-adjoint matrix $\mathbf{H}$ such that $\mathbf{H}^{-1}\mathbf{A}\mathbf{H}$ is normal. By earlier results, we know that all normal matrices are unitarily diagonalizable, i.e.,

$$\mathbf{H}^{-1}\mathbf{A}\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{U}^*,$$

where $\mathbf{D}$ is a diagonal matrix and $\mathbf{U}$ is a unitary matrix. Then, we have that

$$\begin{aligned} \mathbf{H}^{-1}\mathbf{A}\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{U}^* &\implies \mathbf{A} = \mathbf{H}\mathbf{U}\mathbf{D}\mathbf{U}^*\mathbf{H}^{-1} \\ &\implies \mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}, \end{aligned}$$

where $\mathbf{P} = \mathbf{H}\mathbf{U}$. Thus, we see that $\mathbf{A}$ is similar to a diagonal matrix, i.e., we see that $\mathbf{A}$ is diagonalizable.