Numerical analysis: Homework 3

Student: Pratyush Sudhakar (ps2245)

Question 1:

Assume that we are given $A \in \mathbb{R}^{n\times n}$, $A = A^T$, and A has eigenvalue and vector pairs $\{(v_i, \lambda_i)\}_{i=1}^n$. Furthermore, assume that $|\lambda_1| = |\lambda_2| > |\lambda_3| \geq |\lambda_4| \geq \cdots$.

(a) Prove that for any initial guess $v^{(0)}$ such that $v^{(0)}$ is not simultaneously orthogonal to both v_1 and v_2 the power method yields iterates $v^{(k)}$ that converge to lie in the span of v_1 and v_2.

Solution: We know that v_0 is not orthogonal to both v_1 and v_2. Thus, we can write

$$v_0 = c_1v_1 + c_2v_2 + \cdots + c_nv_n,$$

where not both c_1 and c_2 are zero.

Now, by the power method, we have

$$\begin{aligned}
v^{(k)} &= A^k v^{(0)} \\
&= A^k(c_1v_1 + c_2v_2 + \cdots + c_nv_n) \\
&= c_1\lambda_1^k v_1 + c_2\lambda_2^k v_2 + \cdots + c_n\lambda_n^k v_n. \\
(\because A^k v_i &= \lambda_i^k v_i)
\end{aligned}$$

We can write this as,

$$\begin{aligned}
v^{(k)} &= c_1\lambda_1^k \left(v_1 + \frac{c_2}{c_1}\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + \frac{c_n}{c_1}\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n \right); \\
\implies v^{(k)} &= c_1\lambda_1^k \left(v_1 + \frac{c_2}{c_1}\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \sum_{i=3}^n \frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k v_i \right).
\end{aligned}$$

Since, $|\lambda_1| = |\lambda_2| > |\lambda_3| \geq |\lambda_4| \geq \cdots$, we have $\left(\frac{\lambda_2}{\lambda_1}\right) = \pm 1$ and $\left(\frac{\lambda_i}{\lambda_1}\right)^k \to 0$ as $k \to \infty$. Thus, we have

$$\begin{aligned}
v^{(k)} &= c_1\lambda_1^k \left(v_1 + \frac{c_2}{c_1}\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \sum_{i=3}^n \frac{c_i}{c_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k v_i \right) \\
&\approx c_1\lambda_1^k \left(v_1 \pm \frac{c_2}{c_1} v_2 \right); \;\; (\text{as } k \to \infty) \\
&= c_1\lambda_1^k \cdot v_1 \pm c_2\lambda_1^k \cdot v_2;
\end{aligned}$$

where the $\pm$ sign depends on the sign of $\left(\frac{\lambda_2}{\lambda_1}\right)^k$. In both cases, we see that $v^{(k)}$ converges to the span of v_1 and v_2.

(b) What is the rate of convergence of

$$\left(1 - \left\| \left(v^{(k)}\right)^T \begin{bmatrix} v_1 & v_2 \end{bmatrix} \right\|_2^2 \right)^{1/2} ?$$

Solution: We know that $v^{(k)}$ converges to the span of v_1 and v_2. Thus, we can write $v^{(k)}$ as

$$v^{(k)} = \alpha v_1 + \beta v_2 + \sum_{i=3}^{n} \gamma_i v_i,$$

where α and β are non-zero and γ_i are such that $\sum_{i=3}^{n} \gamma_i^2 = 1$.

Now, we have

$$\begin{aligned}\left(v^{(k)}\right)^T \begin{bmatrix} v_1 & v_2 \end{bmatrix} &= \begin{bmatrix} \alpha & \beta \end{bmatrix} \begin{bmatrix} v_1^T v_1 & v_1^T v_2 \\ v_2^T v_1 & v_2^T v_2 \end{bmatrix} \\ &= \begin{bmatrix} \alpha & \beta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \alpha & \beta \end{bmatrix}.\end{aligned}$$

Thus, we have

$$\begin{aligned}\left\| \left(v^{(k)}\right)^T \begin{bmatrix} v_1 & v_2 \end{bmatrix} \right\|_2^2 &= \left\| \begin{bmatrix} \alpha & \beta \end{bmatrix} \right\|_2^2 \\ &= \alpha^2 + \beta^2.\end{aligned}$$

Thus, we have

$$\left(1 - \left\| \left(v^{(k)}\right)^T \begin{bmatrix} v_1 & v_2 \end{bmatrix} \right\|_2^2 \right)^{1/2} = \left(1 - \alpha^2 - \beta^2\right)^{1/2}.$$

Thus, the rate of convergence of $\left(1 - \left\| \left(v^{(k)}\right)^T \begin{bmatrix} v_1 & v_2 \end{bmatrix} \right\|_2^2 \right)^{1/2}$ is $\left(1 - \alpha^2 - \beta^2\right)^{1/2}$.

(c) Does the associated eigenvalue estimate via the Rayleigh quotient necessarily converge in this setting? what about if $\lambda_1 = \lambda_2$?

Question 2:

Assume that we are given $A \in \mathbb{R}^{n \times n}$, $A = A^T$, and A has eigenvalue and vector pairs $\{(v_i, \lambda_i)\}_{i=1}^n$. Furthermore, assume that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell > \lambda_{\ell+1} \geq \cdots \geq \lambda_n \geq 0$. Now, say we run simultaneous iteration (also known an orthogonal iteration) to compute $\operatorname{span}\{v_1, \ldots, v_\ell\}$ and the associated eigenvalues $\lambda_1, \ldots, \lambda_\ell$.

(a) If we denote $\hat{\lambda}_1^{(k)}$ as our guess for λ_1 at iteration k, show that $\lambda_1 \geq \hat{\lambda}_1^{(k)}$ for all k. I.e., our guess for λ_1 converges from below.

(b) As we discussed in class, one reason to discuss convergence of the entire subspace is that it is insensitive to gaps (or the lack thereof) between the first ℓ eigenvalues. If we instead assume $\lambda_1 > \lambda_2 > \cdots > \lambda_\ell > \lambda_{\ell+1} \geq \cdots \geq \lambda_n \geq 0$, would we expect the columns of $V^{(k)}$ (the ON basis for our guess at the invariant subspace of interest at iteration k) to converge to individual eigenvectors (in an appropriate sense)? If so, what might be expect the asymptotic rates of convergence to be? (For this last part a convincing argument suffices, we do not need a formal proof.)

Question 3 (a more challenging, ungraded problem):

Let A be a $n \times n$ matrix that is not diagonalizable, and whose eigenvalue of largest magnitude, denoted λ_1, is associated with a Jordan block of size two. You may assume the rest of the eigenvalues $(\lambda_2, \ldots, \lambda_{n-1})$ are simple. This means that there exists a matrix X such that

$$X^{-1}AX = \begin{bmatrix} \lambda_1 & 1 & \\ & \lambda_1 & \\ & & \Lambda \end{bmatrix}$$

where Λ is a diagonal matrix and $\|\Lambda\|_2 < |\lambda_1|$. Given essentially any initial guess, what, if anything, does the power method applied to A converge to? If it does converge, at what rate does it do so?