Exercise 1

(a) Following our setup $u = v + td$ we note (using the orthogonality of v and d and the fact that they are of unit norm)

$$\begin{aligned}\langle u|u\rangle &= \langle v+td|v+td\rangle \\ &= \langle v|v\rangle + t\langle v|d\rangle + t\langle d|v\rangle + t^2\langle d|d\rangle \\ &= 1+t^2\end{aligned}$$

Therefore

$$\frac{1}{\langle u|u\rangle} = 1 - t^2 + O(t^4).$$

On the other hand

$$\begin{aligned}\langle u|Au\rangle &= \lambda + t\langle v|Ad\rangle + t\langle d|Av\rangle + t^2\langle d|Ad\rangle \\ &= \lambda + t^2\langle d|Ad\rangle,\end{aligned}$$

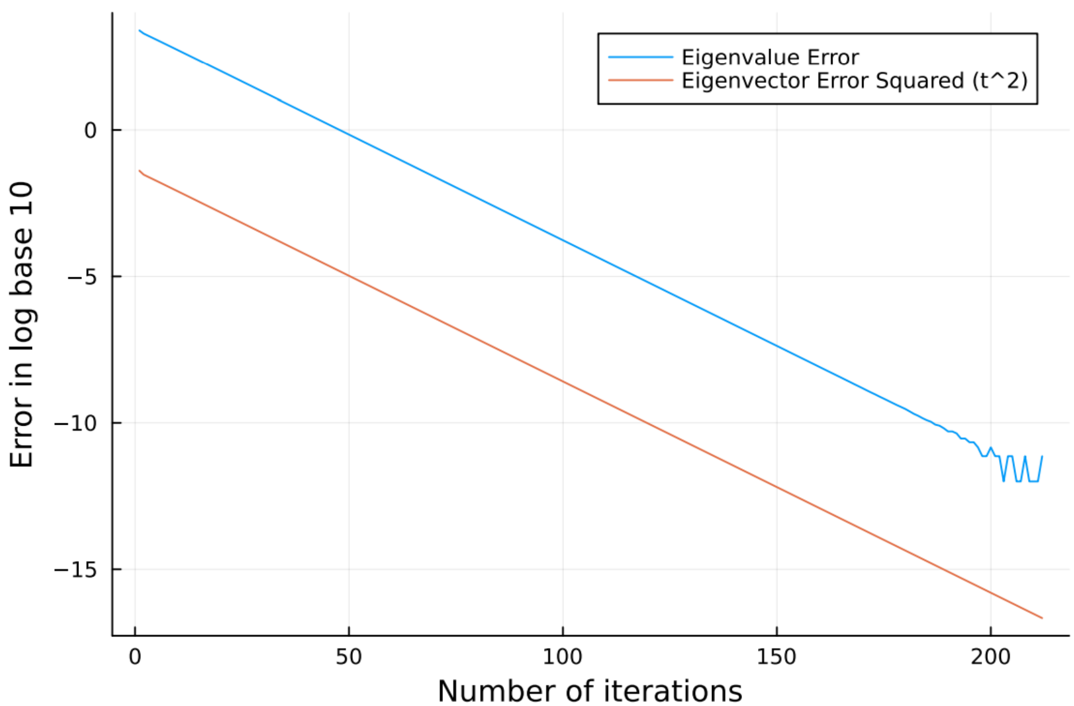
such that overall

$$R_A(u) = (1-t^2)(\lambda + t^2\langle d|Ad\rangle) + O(t^4) = \lambda + t^2(\langle d|Ad\rangle - \lambda) + O(t^4)$$

(b)

$$\lambda - R_A(u) = \varepsilon^2\left(\langle d|Ad\rangle - \lambda\right) + O(t^4)$$

(c)



Exercise 2

(a) For all $x, y \in \mathbb{C}^n$ we have

$$\begin{aligned} 0 &= \langle (x+y), S(x+y) \rangle \\ &= \langle x, Sx \rangle + \langle x, Sy \rangle + \langle y, Sx \rangle + \langle y, Sy \rangle \\ &= \langle x, Sy \rangle + \langle y, Sx \rangle. \end{aligned}$$

(b) Note that $\forall x \in \mathbb{C}^n$

$$\begin{aligned} \langle x, (A - A^H)y \rangle &= \overline{\langle (A - A^H)y, x \rangle} \\ &= \overline{\langle y, (A - A^H)^H x \rangle} \\ &= \overline{-\langle y, (A - A^H)x \rangle}. \end{aligned}$$

(c) First note $\forall x \in \mathbb{C}^n$

$$\langle x, Ax \rangle = \overline{\langle x, Ax \rangle} = \langle Ax, x \rangle = \langle x, A^H x \rangle$$

From this we deduce $\langle x, (A - A^H)x \rangle = 0$ for all $x \in \mathbb{C}^n$. Employing (a) we obtain for arbitrary $x, y \in \mathbb{C}^n$

$$0 = \langle x, (A - A^H)y \rangle + \langle y, (A - A^H)x \rangle \qquad \forall x, y \in \mathbb{C}^n.$$

This gives us

$$\begin{aligned} \overline{\langle y, (A - A^H)x \rangle} &= \overline{-\langle x, (A - A^H)y \rangle} \\ &= \langle y, (A - A^H)x \rangle. \end{aligned}$$

where we have used property (a) in the first and property (b) in the second step. This shows that

$$\langle y, (A - A^H)x \rangle \in \mathbb{R} \qquad \forall x, y \in \mathbb{C}^n.$$

This gives us the freedom to also replace $y \to iy$ and observe

$$\langle iy, (A - A^H)x \rangle = i \langle y, (A - A^H)x \rangle \in \mathbb{R} \qquad \forall x, y \in \mathbb{C}^n.$$

We deduce that

$$\langle y, (A - A^H)x \rangle = 0 \qquad \forall x, y \in \mathbb{C}^n.$$

From this we conclude

$$\langle x, Ay \rangle = \langle x, A^H y \rangle \qquad \forall x, y \in \mathbb{C}^n$$

or $A = A^H$ is Hermitian.

(d)

Let A be a matrix such that $R_A(x)$ is real everywhere. Since $R_A(x) = \langle x, Ax \rangle / \langle x, x \rangle$ and $\langle x, x \rangle \in \mathbb{R}$, we must have $\langle x, Ax \rangle \in \mathbb{R}$ and thus by (c) A must be Hermitian.

Exercise 3

(a)

For any diagonal matrix $D = \text{diag}(d_1, d_2, \ldots, d_n)$ we have

$$\|D\|_F = \sum_{i=1}^{n} |d_i|^2$$

$$\|D\|_p = \max_{v \in \mathbb{C}^n, \|v\|_p = 1} \sqrt[p]{\sum_{i=1}^{n} |d_i|^p |v_i|^p} = \max_{i=1,\ldots,n} |d_i|$$

Since this shows that for a diagonal matrix the Frobenius norm does not agree with any of the matrix norms induced by a p-norm, it cannot be associated to any of the vector p norms.

(b)

For any matrix A we can consider the matrix $A^H A$ which is Hermitian (and positive semidefinite) and thus has an eigendecomposition with nonnegative eigenvalues μ_i

$$A^H A = \sum_{i=1}^{n} \mu_i u_i u_i^H.$$

Let $\mu_{\max}$ denote the largest eigenvalue. We then have

$$\|A\|_2^2 = \mu_{\max} \leq \sum_i \mu_i = \text{tr}(A^H A) = \|A\|_F^2$$

For the case that we have a Hermitian matrix B, it has n real eigenvalues that we can order as

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n.$$

Starting from the Courant-Fisher corollary and using above argument we thus get

$$R_B(x) \leq \max_{0 \neq x \in \mathbb{C}^n} R_B(x) = \lambda_n = \max_i \lambda_i \leq \max_i |\lambda_i| = \|B\|_2 \leq \|B\|_F$$

as desired.

(c)

It is a crude bound since we have the sequence from (b)

$$R_B(x) \leq \lambda_n \leq \|B\|_2 \leq \|B\|_F,$$

and each of these three inequalities can easily have a large gap:

- $R_B(x) \ll \lambda_n$: For example, when the minimum eigenvalue of B is much smaller than the maximum eigenvalue (and corresponding directions x are considered)
- $\lambda_n \ll \|B\|_2$: This happens only if the largest eigenvalue is negative (i.e. the full spectrum is negative and B is negative definite) and the right hand side discards the sign. (Otherwise, for $\lambda_n \geq 0$ we have equality $\lambda_n = \|B\|_2$)
- $\|B\|_2 \ll \|B\|_F$: Whenever there are multiple eigenvalues that have large absolute value. Let us denote $\lambda_{\text{maxabs}} = \max_i |\lambda_i|$ and we get

  $$\|B\|_2^2 = \lambda_{\text{maxabs}}^2 \ll \lambda_{\text{maxabs}}^2 + \sum_{\lambda_i \neq \lambda_{\text{maxabs}}} \lambda_i^2 = \|B\|_F^2.$$

  As an aside, from the same argument we also see an upper bound $\|B\|_F \leq \sqrt{n}\|B\|_2$ which is also true for any matrix, not necessarily Hermitian (e.g. by an analogous proof using singular values).