1. Find an upper bound for the condition number for eigenvector r_j of a non-symmetric matrix A assuming that all its eigenvalues are distinct. In what case will this condition number be large?

Solution: As with the case when A is symmetric, we start with the identity

$$\dot{A}r_j + A\dot{r}_j = \dot{\lambda}_j r_j + \lambda_j \dot{r}_j$$

and expand the perturbation $\dot{r}_j$ in terms of the right eigenvectors $\{r_1, \ldots, r_n\}$ to obtain

$$\dot{A}r_j + A\sum_{i\neq j} m_{ji} r_i = \dot{\lambda}_j r_j + \lambda_j \sum_{i\neq j} m_{ji} r_i.$$

Further, multiplying on the left by ℓ_k yields

$$\ell_k \dot{A} r_j + \ell_k A \sum_{i\neq j} m_{ji} r_i = \dot{\lambda}_j \ell_k r_j + \lambda_j \ell_k \sum_{i\neq j} m_{ji} r_i$$

and as a left eigenvector, $\ell_k A = \lambda_k \ell_k$, so

$$\ell_k \dot{A} r_j + \lambda_k \ell_k \sum_{i\neq j} m_{ji} r_i = \dot{\lambda}_j \ell_k r_j + \lambda_j \ell_k \sum_{i\neq j} m_{ji} r_i$$

Then due to the fact that $L = R^{-1}$, so $\ell_i r_j = \delta_{ij}$ and by assuming $k \neq j$,

$$\ell_k \dot{A} r_j + \lambda_k m_{jk} = \lambda_j m_{jk}$$

and hence solving for m_{jk} yields

$$m_{jk} = \frac{\ell_k \dot{A} r_j}{\lambda_j - \lambda_k}.$$

We can then write

$$\dot{r}_j = \sum_{k\neq j} m_{jk} r_k = \sum_{k\neq j} \frac{\ell_k \dot{A} r_j}{\lambda_j - \lambda_k} r_k$$

and so multiplying by Δt and applying the identities $\Delta r_j = \dot{r}_j \Delta t + O(\|\Delta A\|^2)$ and $\Delta A = \dot{A}\Delta t + O(\|\Delta A\|^2)$,

$$\Delta r_j = \sum_{k\neq j} \frac{\ell_k \Delta A r_j}{\lambda_j - \lambda_k} r_k + O(\|\Delta A\|^2).$$

Now ignoring the $O(\|\Delta A\|^2)$ term and using the known bound for $|\ell_k \Delta A r_j|$,

$$\|\Delta r_j\| = \left\| \sum_{k\neq j} \frac{\ell_k \Delta A r_j}{\lambda_j - \lambda_k} r_k \right\| \leq \sum_{k\neq j} \left\| \frac{\ell_k \Delta A r_j}{\lambda_j - \lambda_k} r_k \right\| = \sum_{k\neq j} \frac{|\ell_k \Delta A r_j|}{|\lambda_j - \lambda_k|} \|r_k\|$$

$$\leq \sum_{k\neq j} \frac{\|\ell_k\| \|\Delta A\| \|r_j\|}{|\lambda_j - \lambda_k|} \|r_k\| = \|\Delta A\| \sum_{k\neq j} \frac{\|\ell_k\| \|r_k\| \|r_j\|}{|\lambda_j - \lambda_k|}.$$

Hence,

$$\kappa(r_j; A) = \lim_{\epsilon\to 0} \max_{\|\Delta A\|=\epsilon} \frac{\|\Delta r_j\|\|A\|}{\|r_j\|\|\Delta A\|} \leq \frac{\|\Delta A\| \sum_{k\neq j} \frac{\|\ell_k\|\|r_k\|\|r_j\|}{|\lambda_j-\lambda_k|}\|A\|}{\|r_j\|\|\Delta A\|}$$
$$= \sum_{k\neq j} \frac{\|\ell_k\|\|r_k\|}{|\lambda_j - \lambda_k|}\|A\| = \|A\| \sum_{k\neq j}(\lambda_j - \lambda_k)^{-1}\|\ell_k\|\|r_k\|$$

which we can further simplify via the bound $\|\ell_k\|\|r_k\| \leq \|R^{-1}\|\|R\|$ to see

$$\kappa(r_j; A) \leq \|A\|\|R^{-1}\|\|R\| \sum_{k\neq j}(\lambda_j - \lambda_k)^{-1}$$

This bound shows that the condition number could be large if there is an eigenvalue close to λ_j, or if the matrix of eigenvectors is ill-conditioned, or if the largest singular value of A is very large (and hence $\|A\|$ is large).

2. Let A be an $n \times n$ matrix. The Rayleigh quotient $Q(x)$ is the following function defined on all $x \in \mathbb{R}^n$:

$$Q(x) \coloneqq \frac{x^\top Ax}{x^\top x}.$$

(a) Let A be symmetric. Prove that $\nabla Q(x) = 0$ if and only if x is an eigenvector of A.

Solution: By the quotient rule, we have that

$$\frac{\partial Q}{\partial x_i} = \frac{(x^\top x)\frac{\partial}{\partial x_i}[x^\top Ax] - (x^\top Ax)\frac{\partial}{\partial x_i}[x^\top x]}{(x^\top x)^2} = \frac{\frac{\partial}{\partial x_i}[x^\top Ax]}{x^\top x} - \frac{(x^\top Ax)\frac{\partial}{\partial x_i}[x^\top x]}{(x^\top x)^2}$$
$$= \frac{\frac{\partial}{\partial x_i}[x^\top Ax]}{x^\top x} - \frac{Q(x)\frac{\partial}{\partial x_i}[x^\top x]}{x^\top x} = \frac{1}{x^\top x}\left(\frac{\partial}{\partial x_i}[x^\top Ax] - Q(x)\frac{\partial}{\partial x_i}[x^\top x]\right).$$

Then towards simplifying this expression further, note that

$$\frac{\partial}{\partial x_i}[x^\top x] = \frac{\partial}{\partial x_i}\sum_{j=1}^n x_j^2 = \sum_{j=1}^n \frac{\partial}{\partial x_i}x_j^2 = \frac{\partial}{\partial x_i}x_i^2 = 2x_i$$

and also that by the product rule,

$$\frac{\partial}{\partial x_i}[x^\top Ax] = \frac{\partial}{\partial x_i}\sum_{j=1}^n x_j(Ax)_j = \sum_{j=1}^n \frac{\partial}{\partial x_i}[x_j(Ax)_j] = \sum_{j=1}^n (Ax)_j\frac{\partial}{\partial x_i}[x_j] + x_j\frac{\partial}{\partial x_i}[(Ax)_j]$$
$$= \sum_{j=1}^n (Ax)_j\frac{\partial}{\partial x_i}[x_j] + \sum_{j=1}^n x_j\frac{\partial}{\partial x_i}[(Ax)_j] = (Ax)_i + \sum_{j=1}^n x_j\frac{\partial}{\partial x_i}\left[\sum_{k=1}^n a_{jk}x_k\right]$$
$$= (Ax)_i + \sum_{j=1}^n x_j\sum_{k=1}^n a_{jk}\frac{\partial}{\partial x_i}x_k = (Ax)_i + \sum_{j=1}^n a_{ji}x_j = (Ax)_i + (A^\top x)_i.$$

Hence, without yet assuming A is symmetric,

$$\frac{\partial Q}{\partial x_i} = \frac{1}{x^\top x}\Big((Ax)_i + (A^\top x)_i - 2Q(x)x_i\Big)$$

and so

$$\nabla Q(x) = \frac{1}{x^\top x}\Big(Ax + A^\top x - 2Q(x)x\Big) = \frac{1}{x^\top x}\Big((A+A^\top)x - 2Q(x)x\Big). \qquad (1)$$

Then if A is symmetric, meaning that $A + A^\top = 2A$,

$$\nabla Q(x) = \frac{1}{x^\top x}\left(2Ax - 2Q(x)x\right) = \frac{2}{x^\top x}\left(Ax - Q(x)x\right).$$

Now if $\nabla Q(x) = 0$ then it must be that $Ax - Q(x)x = 0$, and so

$$Ax = Q(x)x$$

meaning x is an eigenvector of A. Similarly, if x is an eigenvector of A corresponding to λ,

$$Q(x) = \frac{x^\top Ax}{x^\top x} = \frac{x^\top \lambda x}{x^\top x} = \lambda\frac{x^\top x}{x^\top x} = \lambda$$

and so $\nabla Q(x) = 0$. Hence, $\nabla Q(x) = 0$ if and only if x is an eigenvector of A.

(b) Let A be asymmetric. What are the vectors x at which $\nabla Q = 0$?

Solution: Recall from (1) that

$$\nabla Q(x) = \frac{1}{x^\top x}\Big((A+A^\top)x - 2Q(x)x\Big)$$

hence if $\nabla Q(x) = 0$ then $(A + A^\top)x - 2Q(x)x = 0$ and so

$$(A + A^\top)x = 2Q(x)x,$$

meaning $\nabla Q(x) = 0$ when x is an eigenvector of $A + A^\top$.

3. Consider the Rayleigh Quotient Iteration, a very efficient algorithm for finding an eigenpair of a given matrix

```
Input: x_0 ≠ 0 is the initial guess for an eigenvector
v = x_0/||x_0||
for k = 0, 1, 2, ...
   μ_k = v^T A v
   Solve (A − μ_k I)w = v for w
   v = w/||w||.
end for
```

Here is a Matlab program implementing the Rayleigh Quotient Iteration for finding an eigenpair of a random $n \times n$ symmetric matrix starting from a random initial guess:

```
function RayleighQuotient()
n = 100;
A = rand(n);
A = A' + A;
v = rand(n, 1);
v = v / norm(v);
k = 1;
mu(k) = v' * A * v;
tol = 1e-12;
I = eye(n);
res = abs(norm(A * v - mu(k) * v) / mu(k));
fprintf('k = %d: lam = %d\tres = %d\n', k, mu(k), res);
while res > tol
  w = (A - mu(k) * I) \ v;
  k = k + 1;
  v = w / norm(w);
  mu(k) = v' * A * v;
  res = abs(norm(A * v - mu(k) * v) / mu(k));
  fprintf('k = %d: lam = %d\tres = %d\n', k mu(k), res);
end
end
```

(a) Let A be a symmetric matrix with all distinct eigenvalues. Let μ be not an eigenvalue of A. Show that if (λ, v) is an eigenpair of A then $((\lambda-\mu)^{-1}, v)$ is an eigenpair of $(A-\mu I)^{-1}$.

Solution: If (λ, v) is an eigenpair of A, then $Av = \lambda v$. We want to show that

$$(A - \mu I)^{-1} v = (\lambda - \mu)^{-1} v,$$

which by left multiplying both sides by $(A - \mu I)$ is equivalent to

$$\begin{aligned} v = (\lambda - \mu)^{-1}(A - \mu I)v &= (\lambda - \mu)^{-1}(Av - \mu I v) \\ &= (\lambda - \mu)^{-1}(\lambda v - \mu v) = (\lambda - \mu)^{-1}(\lambda - \mu)v = v \end{aligned}$$

which is clearly true. Hence, $(A - \mu I)^{-1} v = (\lambda - \mu)^{-1} v$ is also true, so $((\lambda - \mu)^{-1}, v)$ is an eigenpair of $(A - \mu I)^{-1}$.

(b) The Rayleigh Quotient iteration involves solving the system $(A - \mu_k I)w = v$ for w. The matrix $(A - \mu_k I)$ is close to singular. Nevertheless, this problem is well-conditioned (in exact arithmetic). Explain this phenomenon. Proceed as follows. Without the loss of generality assume that v is an approximation for the eigenvector v_1 of A, and μ is an approximation to the corresponding eigenvalue λ_1. Let $\|v\| = 1$. Write v as

$$v = \left(1 - \sum_{i=2}^{n} \delta_i^2\right)^{1/2} v_1 + \sum_{i=2}^{n} \delta_i v_i,$$

where δ_i, $i = 2, \ldots, n$, are small. Show that the condition number $\kappa((A - \mu I)^{-1}, v)$ (see page 88 in Bindel and Goodman, Principles of scientific computing) is approximately $(1 - \sum_{i=2}^{n} \delta_i^2)^{-1/2}$ which is close to 1 provided that δ_i are small.

Solution: From page 88 of Bindel, we have that the condition number for solving the linear system $Au = b$ is given by

$$\kappa(A^{-1}, b) = \|A^{-1}\| \frac{\|b\|}{\|A^{-1}b\|}$$

and so given the system $(A - \mu_k I)w = v$,

$$\kappa((A - \mu I)^{-1}, v) = \|(A - \mu I)^{-1}\| \frac{\|v\|}{\|(A - \mu I)^{-1}v\|}$$

and so assuming $\|v\| = 1$,

$$\kappa((A - \mu I)^{-1}, v) = \|(A - \mu I)^{-1}\| \frac{1}{\|(A - \mu I)^{-1}v\|}. \tag{2}$$

By part (a), the eigenvalues of $(A - \mu I)^{-1}$ are $(\lambda_i - \mu)^{-1}$. Further, since A is symmetric then so $(A - \mu I)^{-1}$, and so by (1) of Theorem 3, its singular values are

$$|(\lambda_i - \mu)^{-1}| = \frac{1}{|\lambda_i - \mu|}$$

the largest of which will be attained by $i = 1$, since μ approximates λ_1. Hence, by (5) of Theorem 3,

$$\|(A - \mu I)^{-1}\| = \frac{1}{|\lambda_1 - \mu|},$$

and so substituting into (2), we obtain

$$\kappa((A - \mu I)^{-1}, v) = \frac{1}{|\lambda_1 - \mu|} \frac{1}{\|(A - \mu I)^{-1}v\|}.$$

Then writing

$$v = \left(1 - \sum_{i=2}^{n} \delta_i^2\right)^{1/2} v_1 + \sum_{i=2}^{n} \delta_i v_i$$

and noting from part (a) that $(A - \mu I)^{-1}v_i = (\lambda_i - \mu)^{-1}v_i$, we have

$$\begin{aligned}\|(A - \mu I)^{-1}v\| &= \left\| \left(1 - \sum_{i=2}^{n} \delta_i^2\right)^{1/2} (A - \mu I)^{-1}v_1 + \sum_{i=2}^{n} \delta_i (A - \mu I)^{-1}v_i \right\| \\ &= \left\| \left(1 - \sum_{i=2}^{n} \delta_i^2\right)^{1/2} (\lambda_1 - \mu)^{-1}v_1 + \sum_{i=2}^{n} \delta_i (\lambda_i - \mu)^{-1}v_i \right\|.\end{aligned}$$

Hence, since the eigenvalues are distinct, the eigenvectors form an orthonormal basis,

$$\begin{aligned}\kappa((A-\mu I)^{-1}, v) &= \frac{1}{|\lambda_1 - \mu|}\frac{1}{\|(A-\mu I)^{-1}v\|} \\ &= \left\| \left(1 - \sum_{i=2}^{n} \delta_i^2\right)^{1/2} v_1 + \sum_{i=2}^{n} \delta_i \frac{\lambda_1 - \mu}{\lambda_i - \mu} v_i \right\|^{-1} \\ &= \left(\left(1 - \sum_{i=2}^{n} \delta_i^2\right) + \sum_{i=2}^{n} \left(\delta_i \frac{\lambda_1 - \mu}{\lambda_i - \mu}\right)^2\right)^{-1/2} \\ &= \left(1 - \sum_{i=2}^{n} \delta_i^2 \left(1 - \left(\frac{\lambda_1 - \mu}{\lambda_i - \mu}\right)^2\right)\right)^{-1/2}\end{aligned}$$

and so from the fact that $\mu \approx \lambda_1$, we find that

$$\kappa((A-\mu I)^{-1}, v) \approx \left(1 - \sum_{i=2}^{n} \delta_i^2\right)^{-1/2}.$$

(c) It is known that the Rayleigh Quotient iteration converges cubically, which means that the error $e_k := |\lambda - \mu_k|$ decays with k so that the limit

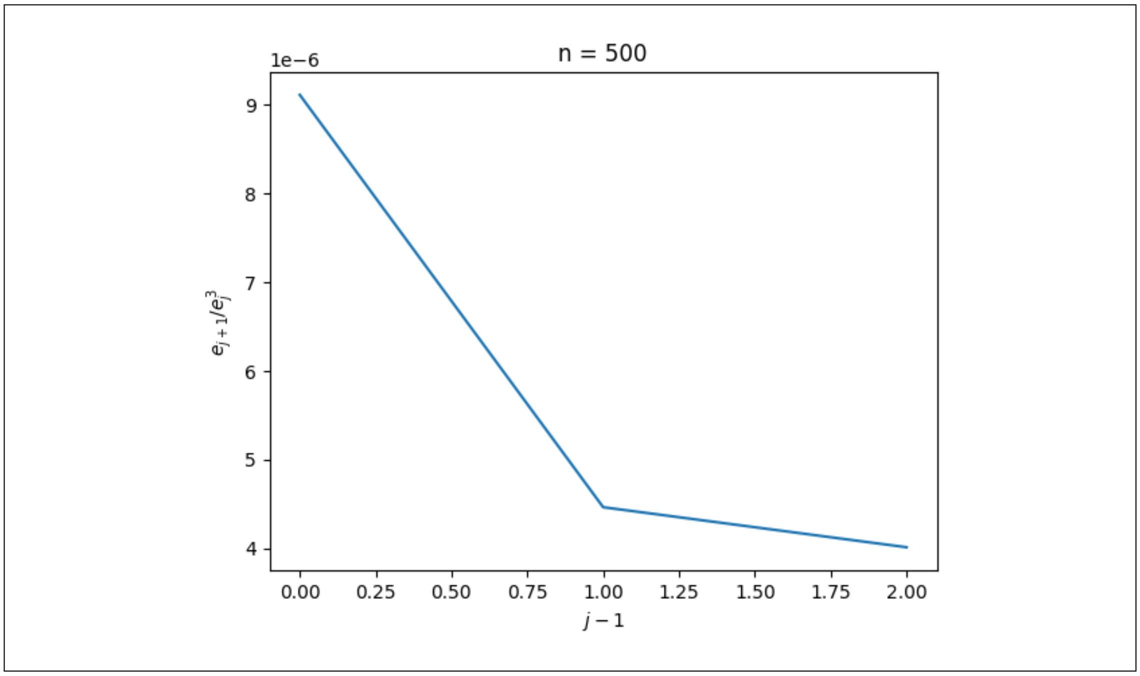
$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k^3} = C \in (0, \infty).$$

This means that the number of correct digits in μ_k triples with each iteration. Try to check this fact experimentally and report your findings. Proceed as follows. Run the program. Treat the final μ_k as the exact eigenvalue. Define $e_j := |\mu_j - \mu_k|$ for $j = 1, \ldots, k-1$. Etc. Pick several values of n and make several runs for each n. Note that you might not observe the cubic rate of convergence due to too few iterations and floating point arithmetic.

Solution: We observe cubic convergence for $n \in \{10, 50, 100, 500, 1000\}$ from the code here. In particular, in each case plotting the ratio

$$\frac{e_{j+1}}{e_j^3}$$

as a function of the iteration number j, the ratio appears to level out given the three available points for each value of n. We show the plot for $n = 500$ below.


n = 500
1e−6
9
8
7
6
5
4
e_{j+1}/e_j^3
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
$j-1$