Theorem 4.4 *Let λ be a simple eigenvalue of A with right eigenvector x and left eigenvector y, normalized so $\|x\|_2 = \|y\|_2 = 1$. Let $\lambda + \delta\lambda$ be corresponding eigenvalue of $A + \delta A$. Then*

$$\delta\lambda = \frac{y^*\delta A x}{y^* x} + O(\|\delta A\|^2)$$

or

$$|\delta\lambda| \le \frac{\|\delta A\|}{|y^* x|} + O(\|\delta A\|^2) = \sec\Theta(y,x)\|\delta A\| + O(\|\delta A\|^2)$$

where $\Theta(y,x)$ is the acute angle between y and x. In other words, $\sec\Theta(y,x) = 1/|y^*x|$ is the condition number *of the eigenvalue λ.*

Proof: Subtract $Ax = \lambda x$ from $(A+\delta A)(x+\delta x) = (\lambda+\delta\lambda)(x+\delta x)$ to get

$$A\delta x + \delta A x + \delta A \delta x = \lambda\delta x + \delta\lambda x + \delta\lambda\delta x\ .$$

Ignore the second order terms (those with two "δ terms" as factors: $\delta A\delta x$ and $\delta\lambda\delta x$) and multiply by y^* to get $y^*A\delta x + y^*\delta A x = y^*\lambda\delta x + y^*\delta\lambda x$.

Now $y^*A\delta x$ cancels $y^*\lambda\delta x$, so we can solve for $\delta\lambda = (y^*\delta A x)/(y^*x)$ as desired. □

Note that a Jordan block has right and left eigenvectors e_1 and e_n, respectively, so the condition number of its eigenvalue is $1/|e_n^*e_1| = 1/0 = \infty$, which agrees with our earlier analysis.

At the other extreme, in the important special case of symmetric matrices, the condition number is 1, so the eigenvalues are always accurately determined by the data:

Corollary 4.1 *Let A be symmetric (or normal: $AA^* = A^*A$). Then $|\delta\lambda| \le \|\delta A\| + O(\|\delta A\|^2)$.*

Proof: If A is symmetric or normal then its eigenvectors are all orthogonal, i.e. $Q^*AQ = \Lambda$ with $QQ^* = I$. So the right eigenvectors x (columns of Q) and left eigenvectors y (conjugate transposes of the rows of Q^*) are identical, and $1/|y^*x| = 1$. □

To see a variety of numerical examples, run the Matlab code referred to in Question 4.14.

Later, in Theorem 5.1, we will prove that in fact $|\delta\lambda| \le \|\delta A\|_2$ if $\delta A = \delta A^T$, no matter how large $\|\delta A\|_2$ is.

Theorem 4.4 is only useful for sufficiently small $\|\delta A\|$. We can remove the $O(\|\delta A\|^2)$ term, and so get a simple theorem true for any size perturbation $\|\delta A\|$, at cost of increasing the condition number by a factor of n:

Theorem 4.5 (Bauer-Fike): *Let A have all simple eigenvalues (i.e. be diagonalizable). Call them λ_i, with right and left eigenvectors x_i and y_i, normalized so $\|x_i\|_2 = \|y_i\|_2 = 1$. Then the eigenvalues of $A + \delta A$ lie in disks B_i where B_i has center λ_i and radius $n\frac{\|\delta A\|_2}{|y_i^* x_i|}$.*

Our proof will use Gershgorin's Theorem (Theorem 2.9 on page 92), which we repeat here:

Gershgorin's Theorem. Let B be an arbitrary matrix. Then the eigenvalues λ of B are located in the union of the n disks defined by $|\lambda - b_{ii}| \le \sum_{j \ne i} |b_{ij}|$, for $i = 1$ to n.

We will also need two simple lemmas:

Lemma 4.1 *Let $S = [x_1, \cdots, x_n]$, the matrix of right eigenvectors. Then*

$$S^{-1} = \begin{bmatrix} y_1^*/y_1^* x_1 \\ y_2^*/y_2^* x_2 \\ \vdots \\ y_n^*/y_n^* x_n \end{bmatrix} .$$

Proof of Lemma: We know $AS = S\Lambda$ where $\Lambda = \text{diag}(\lambda_1 \ldots \lambda_n)$, since the columns x_i of S are eigenvectors. This is equivalent to $S^{-1}A = \Lambda S^{-1}$, so the rows of S^{-1} are conjugate transposes of the left eigenvectors y_i. So $S^{-1} = \begin{bmatrix} y_1^* \cdot c_1 \\ \vdots \\ y_n^* \cdot c_n \end{bmatrix}$ for some constants c_i. But $I = S^{-1}S$ so $1 = (S^{-1}S)_{ii} = y_i^* x_i \cdot c_i$ so $c_i = \frac{1}{y_i^* x_i}$ as desired. □

Lemma 4.2 *If each column of S has 2-norm equal to 1, $\|S\|_2 \le \sqrt{n}$. Similarly, if each row of a matrix has 2-norm equal to 1, its 2-norm is at most $\sqrt{n}$.*

Proof of Lemma: $\|S\|_2 = \|S^T\|_2 = \max_{\|x\|_2=1} \|S^T x\|_2$. Each component of $S^T x$ is bounded by 1 by the Cauchy-Schwartz inequality so $\|S^T x\|_2 \le \left\| [1, \ldots, 1]^T \right\|_2 = \sqrt{n}$. □

Proof of the Bauer-Fike Theorem: We will apply Gershgorin's Theorem to $S^{-1}(A + \delta A)S = \Lambda + F$, where $\Lambda = S^{-1}AS = \text{diag}(\lambda_1, ..., \lambda_n)$ and $F = S^{-1}\delta A S$. The idea is to show that the eigenvalues of $A + \delta A$ lie in balls centered at the λ_i with the given radii. To do this, we take the disks containing the eigenvalues of $\Lambda + F$ that are defined by Gershgorin's Theorem

$$|\lambda - (\lambda_i + f_{ii})| \le \sum_{j \ne i} |f_{ij}|$$