'Oh, complex numbers. They go in circles'

The following example was given in the first lecture of a course to motivate students. We waited until the sixth lecture. The computations we do below were also delayed to the end of that course. The journal reference is at the bottom.

We needed to have a reasonable understanding of complex eigenvectors and eigenvalues. I would remind you that a rotation matrix happily diagonalizes over the complex numbers with two complex eigenvalues of unit length. Multiplying by such a complex number corresponds to a kind of rotation in the argand plane where the real coordinates would happily oscillate back and forth. If this was the x coordinates of our points and the y corrdinates were similarly treated then we would see a rotation.

Imagine we are given n points $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$, ..., $P_n = (x_n, y_n)$. One could for example choose points at random. Now consider an averaging process where we relace the ith point by the average of P_i and P_{i+1} where indices are taken modulo n so that we replace the nth point by the average of P_n and P_1. Repeat this process many times. Of course the points converge to the centroid of the n points, but if we shift the points so that the centroid is the origin and if we rescale (multiply by an appropriate factor at each averaging process) we get an amazing picture. The n points are now arranged well spaced on an ellipse.

We have

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i = 0 \text{ or just } \sum_{i=1}^{n} P_i = \mathbf{0}$$

This will be preserved by the averaging operation. Consider $n = 5$ and let

$$\mathrm{Avg}_5 = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 1/2 & 0 & 0 & 0 & 1/2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Now we see that

$$\mathrm{Avg}_5 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} (x_1+x_2)/2 \\ (x_2+x_3)/2 \\ (x_3+x_4)/2 \\ (x_4+x_5)/2 \\ (x_5+x_1)/2 \end{bmatrix}$$

It is often more convenient to use vector notation so that with

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}, \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix},$$

the new x coordinates after averaging are $\mathrm{Avg}_5\mathbf{x}$.

We can certainly deal with the x coordinates separately from the y coordinates.

To do the averaging process twice, we compute the new x and y coordinates as $\mathrm{Avg}_5(\mathrm{Avg}_5\mathbf{x}) = \mathrm{Avg}_5^2\mathbf{x}$ and $\mathrm{Avg}_5(\mathrm{Avg}_5\mathbf{y}) = (\mathrm{Avg}_5)^2\mathbf{y}$. Now keep averaging and it becomes clear that we need to understand $(\mathrm{Avg}_5)^n$ as $n \to \infty$.

The journal reference is [EV] given below. Our exposition essentially follows theirs and frequently copies from their paper. We have $\text{Avg}_5 = I + C$ where C is the 5×5 circulant:

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Not surprisingly, this matrix has been studied before (not because of this problem). Note that $C^n = I$. The eigenvalues of the $n \times n$ circulant are $\omega_0, \omega_1, \ldots$

$$\omega_j = \cos\left(\frac{2\pi j}{n}\right) + i \cdot \sin\left(\frac{2\pi j}{n}\right)$$

which are the n *roots of unity* with $(\omega_j)^n = 1$. Thus the eigenvalues of $\frac{1}{2}(I + C)$ are $\frac{1}{2}(1 + \omega_j)$ (see assignment question). Moreover

$$\mathbf{v}_j = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 \\ \omega_j \\ \omega_j^2 \\ \vdots \\ \omega_j^{n-1} \end{bmatrix}$$

are an orthonormal basis of eigenvectors of C and hence of $\frac{1}{2}(I + C)$. Note that $\mathbf{v}_0$ is the vector of 1's normalized by multiplying by $\frac{1}{\sqrt{n}}$. Also we are saying $\mathbf{v}_i^H \mathbf{v}_j = 0$ for $i \neq j$. This may not be immediately obvious and note the matrix of eigenvectors will be unitary so this does not force C to be symmetric, which it isn't.

We use $C^T = C^{-1}$ (recall this is a permutation matrix as in an assignment question) and so $C^H = C^{-1}$. Also $\omega_i \omega_{n-i} = 1$ and $\omega_i^H = \overline{\omega_i} = \omega_{n-i} = \omega_i^{-1}$. Note that $C\mathbf{v}_j = \omega_j \mathbf{v}_j$. We compute

$$\mathbf{v}_i^H (C\mathbf{v}_j) = \mathbf{v}_i^H \omega_j \mathbf{v}_j = \omega_j (\mathbf{v}_i^H \mathbf{v}_j)$$

and then

$$(\mathbf{v}_i^H C)\mathbf{v}_j = (\mathbf{v}_i^H (C^{-1})^H)\mathbf{v}_j = (C^{-1}\mathbf{v}_i)^H \mathbf{v}_j = (\omega_i^{-1}\mathbf{v}_i)^H \mathbf{v}_j = \omega_i \mathbf{v}_i^H \mathbf{v}_j.$$

Then with $\omega_i \neq \omega_j$, we deduce $\mathbf{v}_i^H \mathbf{v}_j = 0$. They are orthogonal.

The authors use a different ordering to emphasize certain issues.

$$\begin{array}{lll} \lambda_1 & = \frac{1}{2}(1 + \omega_0) & = 1 \\ \lambda_2 & = \frac{1}{2}(1 + \omega_1) & = (1 + \cos(\frac{2\pi}{n}) + i \cdot \sin(\frac{2\pi}{n})/2 < 1 \\ \lambda_3 & = \frac{1}{2}(1 + \omega_{n-1}) & = \bar{\lambda}_2 \\ \lambda_4 & = \frac{1}{2}(1 + \omega_2) & = (1 + \cos(\frac{4\pi}{n}) + i \cdot \sin(\frac{4\pi}{n})/2 \\ \lambda_5 & = \frac{1}{2}(1 + \omega_{n-2}) & = \bar{\lambda}_4 \end{array}$$

From this ordering reorder the eigenvectors as $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_5$. (In our case with $n = 5$, $\mathbf{z}_2 = \mathbf{v}_1$ and $\mathbf{z}_3 = \mathbf{v}_{n-1} = \mathbf{v}_4$). This ordering of eigenvalues groups together complex conjugates. And the n eigenvalues are on a circle of radius 1 in the complex plane (argand plane) with center $(.5, 0)$. if we follow the order $1, \omega_1, \omega_2, \ldots$ we traverse the circle in counterclockwise order. The new order given as $\lambda_1, \lambda_2, \ldots$ also pairs up the eigenvalues in complex pairs except for $\mathbf{v}_0$, the eigenvector for eigenvalue 1, which is the vector of 1's). We will use $n = 5$. Even or odd n makes a slight difference (-1 is an eigenvalue for n even) but we will focus on n odd with $n = 5$. We have ordered the eigenvalues so that

$$|\lambda_1| = 1 > |\lambda_2| = |\lambda_3| > |\lambda_4| = |\lambda_5|$$

A picture in the complex plane makes you realize that the roots of unity are on a circle and so with the above ordering we are moving through eigenvalues from left to right (to smaller Re components). We also reorder eigenvectors $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_4$ as $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_4, \mathbf{z}_5, \mathbf{z}_3$ so that

$$\mathrm{Avg}_5 \mathbf{z}_k = \lambda_k \mathbf{z}_k \quad k = 1, 2, 3, 4, 5$$

Now consider a unit vector $\mathbf{w} \in \mathbf{C}^n$ with $\mathbf{w} \cdot \mathbf{1} = 0$ (called centroid zero). The vector $\mathbf{w}$ either stands in for the vector $\mathbf{x}$ of x-coordinates of the n points or $\mathbf{w}$ stands in for the vector $\mathbf{y}$ of y-coordinates of the n points. We shifted the points in our code so the centroid was $\mathbf{0}$ and so $\mathbf{w} \cdot \mathbf{1} = 0$. We chose a centroid of $\mathbf{0}$ so the picture does not drift but in this case it also ensures the eigenvector of largest eigenvalue $\lambda_1 = 1$ has coefficient 0 when we expand $\mathbf{w}$ as a linear combination of eigenvectors. So express $\mathbf{w}$ in terms of our orthonormal basis of eigenvectors

$$\mathbf{w} = \gamma_1 \mathbf{z}_1 + \gamma_2 \mathbf{z}_2 + \cdots + \gamma_n \mathbf{z}_n$$

where we note that $\mathbf{w} \cdot \mathbf{1} = 0$ forces $\gamma_1 = 0$. Given the orthonormal basis and the fact $\mathbf{w}$ is a unit vector, we have

$$|\gamma_2^2| + |\gamma_3^2| + \cdots + |\gamma_n^2| = 1$$

We compute

$$(\mathrm{Avg}_5)^k \mathbf{w} = |\lambda_2|^k \left(\gamma_2 \left(\frac{\lambda_2}{|\lambda_2|} \right)^k \mathbf{z}_2 + \overline{\gamma_2} \left(\frac{\lambda_3}{|\lambda_2|} \right)^k \mathbf{z}_3 + \sum_{j=4}^{5} \gamma_j \left(\frac{\lambda_j}{|\lambda_2|} \right)^k \mathbf{z}_j \right)$$

We now wish to consider what happens as $k \to \infty$. The eigenvectors $\mathbf{z}_2, \mathbf{z}_3$ dominate as $k \to \infty$ and so we ignore the terms $j = 4, 5$. In general with more points only λ_2 and λ_3 continue to matter.

With $\lambda_2, \lambda_3 \notin \mathbf{R}$ we must have the coefficients of $\mathbf{z}_2$ and $\mathbf{z}_3$ be complex conjugates (if not as $k \to \infty$ we will have problems).

Let t_i be the ith coordinate of the eigenvector $\mathbf{v}_2$ for λ_2. Then $\overline{t_i}$ will be the ith coordinate of the eigenvector $\bar{\mathbf{v}}_2 = \mathbf{v}_{n-1}$ for λ_3. Then the ith coordinate of $(\mathrm{Avg}_5)^k \mathbf{w}$ is $(|\lambda_2|^k c\omega_2^k t_i + |\lambda_2|^k \overline{c\omega_2}^k \overline{t_i}$ where we have excluded the other terms which vanish as k grows. To consider the behaviour, we ignore the factor $|\lambda_2|^k$ which corresponds to shrinking over time. Matlab rescales for us anyway! The ith coordinate of $(\mathrm{Avg}_5)^k \mathbf{w}$ behaves as an something like $c' \cos(k\theta) + d)$ and hence the coordinates of a single point traces out an ellipse over time (as k grows) and all the points similarly trace out the same ellipse. The spacing of the points comes from the form of the dominant eigenvectors $\mathbf{z}_2$ and $\mathbf{z}_3$ for λ_2 and λ_3.

Let $x_i(k)$ be the ith entry of $(\mathrm{Avg}_5)^k \mathbf{x}$ and let $y_i(k)$ be the ith entry of $(\mathrm{Avg}_5)^k \mathbf{y}$. We plot $(x_i(k), y_i(k))$ as the points. The complex number $\frac{\lambda_2}{|\lambda_2|}$ is $e^{i\theta}$ for some θ. Thus the term $\left(\frac{\lambda_2}{|\lambda_2|}\right)^k$ is $e^{ik\theta}$. Then there is a choice $(a + bi)$ equal to product of γ_2 and t_i (the ith coordinate of $\mathbf{z}_2$). Then for each k we have that

$$x_i(k) = (a + bi)e^{ik\theta} + (a - bi)e^{-ik\theta} = (a + bi)(\cos(k\theta) + i\sin(k\theta)) + (a - bi)(\cos(k\theta) - i\sin(k\theta))$$

$$= 2a\cos(k\theta) - 2b\sin(k\theta)$$

We can write this as $2\sqrt{a^2+b^2}\cos(k\theta+d)$ where d is chose to satisfy $\frac{\cos(d)}{\sin(d)} = -\frac{a}{b}$. So without checking the constants, we have (for all k!)

$$x_i(k) = c\cos(k\theta+d) \text{ and } y_i(k) = c\cos(k\theta+d')$$

Plotted together as $(x_i(k), y_i(k))$ in xy plane, these yield an ellipse. Moreover as k grows we move by 'θ' around the ellipse which we can see in the Matlab pictures.

For more details see the following paper:

References

[EV] Adam N. Elmachtoub, Charles F. Van Loan, From Random Polygons to Ellipse: An Eigenanalysis, *SIAM Review* **52**(2010), 151-170.