A Property of Caratheodory

Seenu S. Reddi
ReddiSS at aol dot com
10/13/2013

A property of Toeplitz matrices attributed to Caratheodory is the following:

For a semi-positive definite Toeplitz matrix with a unique eigenvalue of zero, the polynomial formed from the elements of the eigenvector corresponding to the zero eigenvalue will have roots on the unit circle.

Here the Toeplitz matrix has the $(i, j)^{th}$ element as a function of $|i - j|$, and as an example a 4 X 4 Toeplitz matrix can be represented as follows:

$$\begin{bmatrix} 1 & c_1 & c_2 & c_3 \\ c_1 & 1 & c_1 & c_2 \\ c_2 & c_1 & 1 & c_1 \\ c_3 & c_2 & c_1 & 1 \end{bmatrix}$$

For notational convenience we indicate the Toeplitz matrix as $[[1, c_1, c_2, c_3]]$ and for this discussion, we assume only real matrices. Now we can illustrate the Caratheodory with a numerical example. For T = [[1.00000 -0.78329 0.51987 -0.87595]], the eigenvalues can be calculated as -0.09709, 0.43785, 0.52564 and 3.13361. Subtracting the smallest eigenvalue from the diagonal elements of the matrix, we obtain after normalizing the daigonals to one, [[1.00000, -0.71397, 0.47386, -0.79843]]. For this matrix, the eigenvalues are 0.00000, 0.48760, 0.56762 and 2.94479. The eigenvector that corresponds to the smallest, unique eigenvalue is [-0.54157, 0.45464, 0.49104, -0.50880]. We form a polynomial corresponding to this eigenvector (referred to as zero eigenpolynomial hereafter) as $Z(x) = -0.54157 + 0.45464 * x + 0.49104 * x^2 - 0.50880 * x^3$. Solving for $Z(x) = 0$, we see the roots are $\exp(i * \pi)$, $\exp(i * 1.495)$ and $\exp(-i * 1.495)$ which lie on the unit circle thus confirming the Caratheodory's property. In this discussion for simplicity we indicate these roots in degrees as 180°, 85.40°, and -85.40°.

Note Caratheodory expressed this property *inter alia* in 1900's when there was no fast electronic computers available and tools like Mathematica, MatLab/SCiLab, etc. were non-existent. Thus it is all more remarkable that such a property can be derived and proved using only mathematical analysis and possible hand calculations. However proofs are long and laborious and do not reflect the underlying structure that leads to such a nice and beautiful result. We approach the problem from a different viewpoint in a numerical setting so as to illustrate why the Toeplitz matrices have this property and understand the underlying structure that leads to this remarkable property. It should be mentioned that one can see intuitively what is going on from the numerical results but unfortunately we do not have at this time simple proofs to accompany the insight.

We define Vandermonde Unit Vectors (VUV) as $[1, x, x^2, x^3 \dots x^{n-1}]$ where $x = \exp(i * \theta)$, $0 \leq \theta \leq 2\pi$. VUVs can span a n-dimensional space by selecting n different θ's because the Vandermonde determinant formed from these vectors will be non-zero if the θ's are different. Now we note that from $VUV(\theta) = [1, \exp(i * \theta), \exp(i * 2\theta), \exp(i * 3\theta) \,..\, \exp(i * (n - 1)\theta)]$, we can form the rows of a Toeplitz matrix by multiplying with $\exp(i * -m\theta)$, $m = 1, .., n - 1$. Thus by multiplying with $\exp(i * -\theta)$, we have the second row:

$$\exp(i * -\theta) * VUV(\theta) = [\exp(i * -\theta), 1, \exp(i * \theta) \,..\, \exp(i * (n - 2)\theta)]$$

and the third row by multiplying with $\exp(i * -2\theta)$: (1)

$$\exp(i * -2\theta) * VUV(\theta) = [\exp(i * -2\theta), \exp(i * -\theta), 1, \exp(i * \theta) \,..\, \exp(i * (n - 3)\theta)]$$

and so on. This observation connects the *Toeplitzness* with VUVs and leads us to represent a Toeplitz matrix using VUVs. Since the VUVs span the n-dimensional space, for a given Toeplitz matrix $[[1, c_1, c_2, c_3, .. c_{n-1}]]$, select any arbitrary basis vectors $VUV(\theta_i)$ with different θ_i and solve for $[\lambda_1, \lambda_2, \lambda_3, .. \lambda_n]$:

$$\lambda_1 * VUV(\theta_1) + \lambda_2 * VUV(\theta_2) + .. + \lambda_n * VUV(\theta_n) = [1, c_1, c_2, c_3, .. c_{n-1}]$$

We indicate column vectors with [..] and row vectors (..). The above equation can be expressed as:

$$[VUV(\theta_1)\ VUV(\theta_2) \,..\, VUV(\theta_n)]\ [\lambda_1, \lambda_2, \lambda_3, .. \lambda_n] = [1, c_1, c_2, c_3, .. c_{n-1}]$$

and solved for λ_i: (2)

$$[\lambda_1, \lambda_2, \lambda_3, .. \lambda_n] = [VUV(\theta_1)\ VUV(\theta_2) \,..\, VUV(\theta_n)]^{-1}\ [1, c_1, c_2, c_3, .. c_{n-1}]$$

Now the *Toeplitzness* and VUV connection expressed by (1) can be invoked to express the other rows of the Toeplitz matrix from (2). For instance the second row = $[c_1, 1, c_1, c_2, c_3, .. c_{n-2}]$ can be expressed in terms of λ_i and θ_i:

$$[[VUV(\theta_1)]\ [VUV(\theta_2)] \,..\, [VUV(\theta_n)]]\ [\lambda_1\exp(i * -\theta_1), \lambda_2\exp(i * -\theta_2), .. \lambda_n\exp(i * -\theta_n)] = [c_1, 1, c_1, c_2, c_3, .. c_{n-2}] \quad (3)$$

By writing (3) for all the rows of the Toeplitz matrix, we can see the matrix can be expressed as a product:

$[[1, c_1, c_2, c_3, .. c_{n-1}]] =$ (4)

$$[[VUV(\theta_1)]\ [VUV(\theta_2)]\ ..\ [VUV(\theta_n)]]\ \Lambda\ [(VUV(-\theta_1))\ (VUV(-\theta_2))\ ..\ (VUV(-\theta_n))]$$

where Λ is the diagonal matrix with diagonal elements λ_1, λ_2, λ_3, .. and λ_n. In fact (4) can be seen by multiplying the matrices but by taking the longer route, one can see the connection between the Toeplitz matrix and the Vandermonde unit vectors.

Now we consider the Toeplitz matrix [[1.00000, -0.71397, 0.47386]] from our example matrix above with the last element removed. Since we are considering only real matrices, θ's are to be selected so that VUVs are either real or form conjugate pairs. We denote a VUV by a value t in degrees between 0 and 180° and indicate by (t) = [1, exp(i * T), .. (exp(i * (n-1)T)] where T = t * π / 180 radians. Since the matrix is real, we can either select {(1), (t), (-t)} or ((-1), (t), (-t)}. Running thru' all possible values for t, we find out that there are only *two* sets of vectors that can represent the Toeplitz matrix. They are {(1), (147.84), (-147.84)} and {(-1), (85.40), (-85.40)}. For a given basis {(t1), (t2), (t3)}, let $\{\lambda_1, \lambda_2, \lambda_3\}$ be the solution so that λ_1* (t1) + λ_2 * (t2) + λ_3 * (t3) = [1.00000, -0.71397, 0.47386]. It turns out only for these two sets, solutions are real and for any other set of basis vectors, solutions are not real and the matrix becomes asymmetric.

First Set:

Basis: {(1), (147.84), (-147.84)}
$\{\lambda_1, \lambda_2, \lambda_3\}$ = {0.0717592 - 2.567D-17i , 0.4641204 - 0.0000817i , 0.4641204 + 0.0000817i}

Toeplitz matrix from Basis and $\{\lambda_1, \lambda_2, \lambda_3\}$:

1. - 5.342D-17i	- 0.7141439 + 5.333D-17i	0.4741545 - 1.487D-16i
- 0.71397 - 5.551D-17i	1. - 5.551D-17i	- 0.7141439 + 1.388D-17i
0.47386 + 1.110D-16i	- 0.71397 - 1.110D-16i	1. - 5.551D-17i

Second Set:

Basis: {(-1), (85.40), (-85.40)}
$\{\lambda_1, \lambda_2, \lambda_3\}$ = {0.7352253 - 2.481D-18i, 0.1323873 - 0.0000104i, 0.1323873 + 0.0000104i}

Toeplitz matrix from Basis and $\{\lambda_1, \lambda_2, \lambda_3\}$:

1. - 2.480D-18i	- 0.7140114 + 2.481D-18i	0.4738534 - 2.480D-18i
- 0.71397	1. - 1.735D-18i	- 0.7140114
0.47386 - 3.469D-18i	- 0.71397 + 2.385D-18i	1. - 3.469D-18i

Wrong Set:

Basis: {(1), (147), (147)}
$\{\lambda_1, \lambda_2, \lambda_3\}$ = { 0.0751327 - 7.302D-18i, 0.4624336 + 0.0123419i, 0.4624336 - 0.0123419i}

Toeplitz matrix from Basis and $\{\lambda_1, \lambda_2, \lambda_3\}$:

1. - 3.469D-17i	-0.6870824 + 7.112D-17i	0.4287603 - 1.292D-16i
- 0.71397 - 5.551D-17i	1.	- 0.6870824 + 8.327D-17i
0.47386 + 1.110D-16i	- 0.71397 - 1.110D-16i	1.

We see when the solutions are real, λ_i's are real and the Toeplitz matrix is reproduced properly but in the wrong set, some λ_i's are not real and the reproduced matrix is not symmetric. Though it can be proved that λ_i's have to be real for the right solution, it appears the proof that only two solutions are possible may not be that simple. However we show later the two solutions are connected to some other factors.

Once we have the basis for the non-singular n X n Toeplitz matrix, it can be extended to (n + 1) X (n + 1) Toeplitz matrix by extending the individual VUV basis vectors to n + 1 elements as follows. A given VUV(θ) = [1, exp(i * θ), exp(i * 2θ), exp(i * 3θ) .. exp(i * (n - 1)θ)] can be extended to n + 1 elements [1, exp(i * θ), exp(i * 2θ), exp(i * 3θ) .. exp(i * (n - 1)θ), exp(i * nθ)] and the Toeplitz matrix $T_n = [1, c_1, c_2, c_3, .. c_{n-1}]$ can be extended to $T_{n+1} = [1, c_1, c_2, c_3, .. c_{n-1,} x]$ by defining x to be:

$$x = \lambda_1 \exp(i * n\theta_1) + \lambda_2 \exp(i * n\theta_2) + .. + \lambda_n \exp(i * n\theta_{n-1}) \qquad (5)$$

Since the extended basis vectors span n-dimensional space (i.e., the matrix composed of these basis vectors is of rank n), T_{n+1} is of rank n and thus have an eigenvalue of zero. Also the eigenvector corresponding to the zero eigenvalue will be in the null subspace and orthogonal to the basis vectors. If $\xi = [\xi 1, \xi 2, \ldots, \xi n]$ is the eigenvector corresponding to the zero eigenvalue, then we have:

$$T_{n+1} \xi = 0 \qquad (6)$$

From (4) we can conclude that ξ is orthogonal to (VUV($-\theta_1$)), (VUV($-\theta_2$)), .., and (VUV($-\theta_n$)). The polynomial formed from ξ has roots on the unit circle namely exp(i * $-\theta_1$), exp(i * $-\theta_2$), .., exp(i * $-\theta_{n-1}$) thus confirming the Caratheodory's property.

Taking our numerical example above, we have from the first set:

Basis: {(1), (147.84), (-147.84)}
$\{\lambda_1, \lambda_2, \lambda_3\}$ = {0.0717592 - 2.567D-17i , 0.4641204 - 0.0000817i , 0.4641204 + 0.0000817i}

Extending the matrix to a 4 X 4 matrix:

$\lambda_1 * (1)_4 + \lambda_2 * (147.84)_4 + \lambda_3 * (-147.84)_4 = (1.00000, -0.71397, 0.47386, 0.17680)$

And for the second set:

Basis: {(-1), (85.40), (-85.40)}
$\{\lambda_1, \lambda_2, \lambda_3\}$ = {0.7352253 - 2.481D-18i, 0.1323873 - 0.0000104i, 0.1323873 + 0.0000104i}

$\lambda_1 * (1)_4 + \lambda_2 * (85.40)_4 + \lambda_3 * (-85.40)_4 = (1.00000, -0.71397, 0.47386, -0.79843)$

As we may recall, the first Toeplitz matrix we considered as an example is [[1.00000, -0.71397, 0.47386, -0.79843]]. And now we see it has a dual [[1.00000, -0.71397, 0.47386, 0.17680]] and both of them have roots on the unit circle for their zero eigenpolynomials.

The two values -0.79843 and 0.17680 can also be gotten by a different argument. Consider the matrix [[1.00000, -0.71397, 0.47386, x]] and determine x so that the determinant of [[1.00000, -0.71397, 0.47386, x]] is zero. The matrix will become semi-definite with unique eigenvalue of zero as its determinant is zero and it becomes singular. Evaluating the determinant we have:

Determinant[[1.00000, -0.71397, 0.47386, x]] = 0
$\Rightarrow 0.0692051 - 0.304756\, x - 0.490247\, x^2 = 0$
$\Rightarrow x = -0.798438$ and 0.1768

We used Mathematica to evaluate the determinant in a symbolic form. One can also evaluate the determinant numerically for various values of x and determine the roots. The following theorem asserts there are two possible values for x that will make the non-singular Toeplitz matrix singular (or semi-definite with a unique eigenvalue of zero).

Theorem: A n X n non-singular Toeplitz matrix $T_n = [[1, c_1, c_2, c_3, .. c_{n-1}]]$ can be extended to a semi-definite Toeplitz matrix $T_{n+1} = [[1, c_1, c_2, c_3, .. c_{n-1}, x]]$ with a unique eigenvalue of zero in two ways. x satisfies the quadratic equation:

$$\text{Determinant}([[1, c_1, c_2, c_3, .. c_{n-1}, x]]) = 0 \quad (7)$$

The determinantal equation in x can be seen to be quadratic by looking at its expansion.

To summarize our approach and what has been proved or not, we list the following items:

1. Any non-singular Toeplitz matrix $[[1, c_1, c_2, c_3, .. c_{n-1}]]$ can represented in only *two* ways using VUV's. When the VUVs are used as basis vectors, the coefficients have to be real to express the vector $[1, c_1, c_2, c_3, .. c_{n-1}]$ for both of these ways. This assertion is unproven and based on numerical evidence.

2. There are only two ways of obtaining a semi-definite matrix [[1, c_1, c_2, c_3, .. $c_{n-1,}$ x]] by extending the VUV's from n elements to n + 1 elements and expressing x as in Equation (5). These are the only possible semi-definite matrices which contains the leading submatrix [[1, c_1, c_2, c_3, .. c_{n-1}]]. Follows from 1.

3. For the two Toeplitz matrices obtained in 2 the Caratheodory's property is satisfied. Thus the Caratheodory's property follows from its dual representation in 1.

4. The determinantal expansion of Determinant[[1, c_1, c_2, c_3, .. $c_{n-1,}$ x]] = 0 gives two roots for x and thus two Toeplitz matrices which are semi-definite. Numerical evidence shows these two matrices to be the same as in 2.

From numerical evidence it appears possible to prove the Caratheodory's property by demonstrating that there are only two possible semi-definite n + 1 X n + 1 Toeplitz matrices for a given leading n X n Toeplitz submatrix [[1, c_1, c_2, c_3, .. c_{n-1}]] and for these two matrices, the zero eigenpolynomial has roots on the unit circle.

An extensive simulation was performed on 5 X 5 real Toeplitz matrices of the form [[1, c_1, c_2, c_3, x]] and it was found that x is always real and has two solutions. When combined with the Caratheodory's property, it can be asserted that this type of matrices always have a dual representation with VUVs. These matters should be researched in the future.

What is interesting is the dual values possible for semi-definiteness (since cara in Caratheodory means face/head in Greek and has the Indo-European root *siras* meaning head in Sanskrit, maybe these dual values should be referred to as head and tail). Usually a harmonic analysis is using the Toeplitz matrix and the Caratheodory's property is used to deduce the harmonic nature as expressed by the roots on the unit circle of the zero eigenpolynomial. As our analysis points out, there are two possible cases and they provide different sets of roots. This may require a careful interpretation of the harmonic analysis.

We conclude with a 11 X 11 Toeplitz matrix as an example. The matrix [[1.00000, 0.36252, -0.04367, 0.00327, 0.19497, -0.22068, 0.06096, 0.16532, 0.02092, -0.28726, 0.19065]] has the eigenvalues:

0.00000, 0.02330, 0.24901, 0.65600, 0.97845, 1.17248, 1.42632, 1.49510, 1.58379, 1.65957, 1.75598

The zero eigenpolynomial has the roots:

28.95º, -28.95º, 104.84º, -104.84º, 88.37º, -88.37º, 58.24º, -58.24º, 180.00º, 0.00º

Determinant[[1.00000, 0.36252, -0.04367, 0.00327, 0.19497, -0.22068, 0.06096, 0.16532, 0.02092, -0.28726, x]] = $0.0000428923 + 0.000643451\, x - 0.00454485\, x^2$

x = -0.0494135, 0.190991

The second root corresponds to our original Toeplitz matrix (within numerical accuracy). The other value of -0.0494135 gives the Toeplitz matrix [[1.00000, 0.36252, -0.04367, 0.00327, 0.19497, -0.22068, 0.06096, 0.16532, 0.02092, -0.28726, -0.0494135]] and its eigenvalues:

0.00000, 0.00379, 0.37539, 0.65181, 0.86729, 1.18001, 1.37567, 1.49640, 1.53916, 1.74983, 1.76064

The zero eigenpolynomial has the roots:

178.18º, -178.18º, 47.74º, -47.74º, 103.40º, -103.40º, 76.63º, -76.63º, 15.38º, -15.38º

The roots can be seen to be different and an interesting issue is whether the Caratheodory's property indicates any harmonicity or just a mathematical curiosity to which harmonic properties are attached.

Acknowledgment: We used SciLab and Mathematica for our simulations. Discussions with Florian Bünger have been helpful in re-visiting the problem.