A NOVEL METHOD TO DERIVE EXPLICIT KLT KERNEL FOR AR(1) PROCESS

*Mustafa U. Torun and Ali N. Akansu**

Department of Electrical and Computer Engineering
New Jersey Institute of Technology
University Heights, Newark, NJ 07102 USA

ABSTRACT

The derivation of explicit Karhunen-Loève transform (KLT) kernel for an auto-regressive order one, AR(1), process has been of a major implementation concern. The main reason is due to difficulties in finding the roots of a transcendental tangent equation as required in the derivation. We propose a novel method to derive explicit KLT kernel for discrete AR(1) process. The merit of the proposed technique is highlighted. The proposed method has a good potential to make real-time KLT applications more feasible in the coming years.

1. INTRODUCTION

The Karhunen-Loève transform (KLT) is the optimal block transform. Its basis functions are generated for a given signal covariance matrix [1]. It is a signal dependent transform. In contrast, the popular transforms like discrete Fourier transform (DFT) and discrete cosine transform (DCT) have their fixed kernel to define an orthogonal set regardless of signal statistics. Therefore, they are called signal independent or fixed transforms. Fast implementation of KLT is of great interest in various disciplines. There were prior studies to derive closed form kernel expressions for certain classes of stochastic processes reported in the literature. In particular, the derivation of such a kernel in its implicit form for processes with exponential correlation was reported [2–6]. Those methods require us to solve a transcendental tangent equation by using either numerical techniques with convergence concerns, or complex methods for explicit expression of KLT kernel. In this paper, we introduce a simple method to derive explicit KLT kernel for discrete AR(1) process by using an efficient root-finding technique introduced in [7]. A summary of this derivation is presented in Section 4 [8]. We show and validate its merit with performance comparisons given in Section 5. Then, we make our concluding remarks in the last section.

*Corresponding author: akansu@njit.edu

2. ORTHOGONAL TRANSFORMS

A family of linearly independent N orthonormal discrete-time sequences, $\{\phi_k(n)\}$, on the interval $0 \leq n \leq N-1$ satisfies the relationship [1]

$$\sum_{n=0}^{N-1} \phi_k(n)\phi_l^*(n) = \delta_{k-l}, \tag{1}$$

where δ_k is the Kronecker delta function. In matrix form, $\{\phi_k(n)\}$ are the rows of the transform matrix, also called basis functions,

$$\Phi = [\phi_k(n)] : k, n = 0, 1, ..., N-1, \tag{2}$$

with the orthogonality property stated as

$$\Phi\Phi^{-1} = \Phi\Phi^{*\mathrm{T}} = \mathbf{I}, \tag{3}$$

where $*\mathrm{T}$ indicates conjugated and transposed version of a matrix and $\mathbf{I}$ is $N \times N$ identity matrix. A signal vector $\mathbf{x} = \begin{bmatrix} x(0) & x(1) & \cdots & x(N-1) \end{bmatrix}^{\mathrm{T}}$ is mapped into the orthonormal space through forward transform operator

$$\boldsymbol{\theta} = \Phi\mathbf{x}, \tag{4}$$

where $\boldsymbol{\theta} = \begin{bmatrix} \theta(0) & \theta(1) & \cdots & \theta(N-1) \end{bmatrix}^{\mathrm{T}}$ is transform coefficients vector. Similarly, the inverse transform yields the signal vector

$$\mathbf{x} = \Phi^{-1}\boldsymbol{\theta}. \tag{5}$$

We assume that the vector $\mathbf{x}$ is populated by a wide-sense stationary (WSS) stochastic process. The correlation and covariance matrices of such a random vector process $\mathbf{x}$ are defined, respectively,

$$\begin{aligned} \mathbf{R}_x &= E\left\{\mathbf{x}\mathbf{x}^{*\mathrm{T}}\right\}, \\ \mathbf{C}_x &= E\left\{(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^{*\mathrm{T}}\right\} = \mathbf{R}_x - \boldsymbol{\mu}\boldsymbol{\mu}^{*\mathrm{T}}, \end{aligned} \tag{6}$$

where $\boldsymbol{\mu} = \begin{bmatrix} \mu & \mu & \cdots & \mu \end{bmatrix}^{\mathrm{T}}$. The covariance matrix of transform coefficients is derived as follows

$$\mathbf{R}_\theta = E\left\{\boldsymbol{\theta}\boldsymbol{\theta}^{*\mathrm{T}}\right\} = E\left\{\Phi\mathbf{x}\mathbf{x}^{*\mathrm{T}}\Phi^{*\mathrm{T}}\right\} = \Phi\mathbf{R}_x\Phi^{*\mathrm{T}}. \tag{7}$$

3. EIGENANALYSIS OF AR(1) PROCESS

Auto-regressive (AR), moving average (MA), and auto-regressive moving average (ARMA) have been popular source models to describe real world signals. AR source model has been successfully used in speech processing for decades. AR model with order one, AR(1), is a first approximation to many natural signals like images, and employed in various applications. AR(1) signal is generated through the first order regression formula written as [1]

$$x(n) = \rho x(n-1) + \xi(n), \tag{8}$$

where $\xi(n)$ is a white noise sequence with zero mean, i.e. $E\{\xi(n)\xi(n+k)\} = \sigma_\xi^2 \delta_{n-k}$. The first order correlation coefficient, ρ, is real in the range of $-1 < \rho < 1$, and the variance of $x(n)$ is $\sigma_x^2 = (1-\rho^2)^{-1}\sigma_\xi^2$. Auto-correlation sequence of $x(n)$ is expressed as

$$R_{xx}(k) = E\{x(n)x(n+k)\} = \sigma_x^2 \rho^{|k|}; k = 0, \pm 1, \pm 2, \ldots. \tag{9}$$

The resulting Toeplitz correlation matrix of size $N \times N$ is defined as

$$\mathbf{R}_x = [R_x(k,n)] = \sigma_x^2 \rho^{|n-k|}; k, n = 0, 1, ..., N-1. \tag{10}$$

From linear algebra, it is known that an eigenvalue λ and an eigenvector $\boldsymbol{\phi}$ with size $N \times 1$ of a matrix $\mathbf{R}_x$ with size $N \times N$ must satisfy the eigenvalue equation [1, 2, 9]

$$\mathbf{R}_x \boldsymbol{\phi} = \lambda \boldsymbol{\phi}. \tag{11}$$

It is rewritten as

$$\mathbf{R}_x \boldsymbol{\phi} - \lambda \mathbf{I} \boldsymbol{\phi} = (\mathbf{R}_x - \lambda \mathbf{I})\, \boldsymbol{\phi} = \mathbf{0}, \tag{12}$$

such that $(\mathbf{R}_x - \lambda \mathbf{I})$ is singular. Namely,

$$\det(\mathbf{R}_x - \lambda \mathbf{I}) = 0. \tag{13}$$

We assume that $\mathbf{R}_x$ is non-defective, i.e. its eigenvectors with different eigenvalues are linearly independent, and this determinant is a polynomial in λ of degree N. Therefore, (13) has N roots and (12) has N solutions for $\boldsymbol{\phi}$ that result in the eigenpair set $\{\lambda_k, \boldsymbol{\phi}_k\}$ where $0 \le k \le N-1$. Hence, we can write the eigendecomposition for a non-defective $\mathbf{R}_x$ with distinct eigenvectors as follows

$$\mathbf{R}_x = \mathbf{A}_{KLT}^{*\mathrm{T}} \Lambda \mathbf{A}_{KLT} = \sum_{k=0}^{N-1} \lambda_k \boldsymbol{\phi}_k \boldsymbol{\phi}_k^{*\mathrm{T}}, \tag{14}$$

where $\boldsymbol{\Lambda} = diag(\lambda_k); k = 0, 1, \ldots, N-1$, and k*th* column of $\mathbf{A}_{KLT}^{*\mathrm{T}}$ matrix is the k*th* eigenvector $\boldsymbol{\phi}_k$ of $\mathbf{R}_x$ with the corresponding eigenvalue λ_k. Note that $\{\lambda_k = \sigma_k^2\}\, \forall k$, for the given $\mathbf{R}_x$ where σ_k^2 is the variance of the k*th* transform coefficient, θ_k. The eigenvalues of $\mathbf{R}_x$ for an AR(1) process defined in (10) are derived to be in the form [3]

$$\sigma_k^2 = \lambda_k = \frac{1-\rho^2}{1 - 2\rho\cos(\omega_k) + \rho^2}; 0 \le k \le N-1, \tag{15}$$

where $\{\omega_k\}$ are the positive roots of the following equation

$$\tan(N\omega) = -\frac{(1-\rho^2)\sin(\omega)}{\cos(\omega) - 2\rho + \rho^2\cos(\omega)}, \tag{16}$$

that is equivalent to [8]

$$\left(\tan\frac{\omega N}{2} + \gamma \tan\frac{\omega}{2}\right)\left(\tan\frac{\omega N}{2} - \frac{1}{\gamma}\cot\frac{\omega}{2}\right) = 0.$$
$$\gamma = (1+\rho)\,/\,(1-\rho)\,, \tag{17}$$

The resulting KLT matrix of size $N \times N$ is expressed with the explicit kernel as [3]

$$\mathbf{A}_{KLT} = [A(k,n)] = c_k \sin\left[\omega_k\left(n - \frac{N-1}{2}\right) + \frac{(k+1)\pi}{2}\right]$$
$$c_k = \left(\frac{2}{N+\lambda_k}\right)^{1/2},\ 0 \le k, n \le N-1. \tag{18}$$

Note that the roots of the transcendental tangent equation in (17), $\{\omega_k\}$, are required in the KLT kernel as defined in (18). There are well-known numerical methods like *secant method* [10] to approximate roots of the tangent equation given in (17) in order to solve it explicitly. We focus on an efficient root finding method proposed by Luck and Stevens [7] to find explicit solutions for transcendental equations including the tangent equation of (17). It leads us to explicit definition and derivation of KLT kernel given in (18) for an AR(1) process. The derivation details of (15) and (16) are given in [8].

4. A NOVEL METHOD FOR EXPLICIT KLT KERNEL OF AR(1) PROCESS

In order to derive an explicit expression for the discrete KLT kernel, according to (17), we need to calculate the first $N/2$ positive roots of the following two transcendental equations

$$\tan\frac{\omega N}{2} = -\gamma \tan\frac{\omega}{2} \tag{19}$$
$$\tan\frac{\omega N}{2} = \frac{1}{\gamma}\cot\frac{\omega}{2}. \tag{20}$$

Note that roots of (19) and (20) correspond to the even and odd indexed eigenvalues and eigenvectors, respectively [8]. Fig. 1 displays functions $\tan(\omega N/2)$ and $-\gamma\tan(\omega/2)$ for $N = 8$ and various values of ρ. It is apparent from the figure

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$N=4$	0.157	0.815	1.584	2.362												
$N=8$	0.109	0.423	0.801	1.188	1.577	1.968	2.359	2.750								
$N=16$	0.075	0.224	0.408	0.599	0.793	0.988	1.183	1.378	1.574	1.770	1.966	2.162	2.358	2.554	2.750	2.946

Table 1. The values of $\{\omega_k\}$ for $\rho = 0.95$ and $N = 4,\ 8,\ 16$.

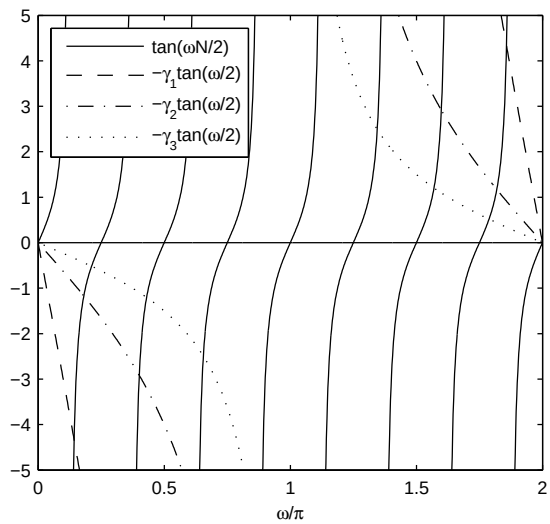

Fig. 1. Functions $\tan(\omega N/2)$ and $-\gamma\tan(\omega/2)$ for $N = 8$ and various values of ρ for $\rho_1 = 0.9$, $\rho_2 = 0.6$, and $\rho_3 = 0.2$ where $\gamma_i = (1+\rho_i)/(1-\rho_i)$, $i = 1, 2, 3$.

that for the m*th* root of (19), a suitable choice for the closed path C discussed in [7] is a circle of radius

$$R_m = \begin{cases} \pi/4N & m \le 2 \\ \pi/2N & m > 2 \end{cases}, \tag{21}$$

centered at $h_m = (m - 1/4)\pi/N$ where $1 \le m \le N/2$. We reconfigure (19) and rather look for the poles of the following inverse function

$$g(\omega) = \frac{1}{\tan(\omega N/2) + \gamma\tan(\omega/2)}. \tag{22}$$

The function $w(\theta)$ defined in (4) of [7] for this case is defined as

$$\begin{aligned} w_m(\theta) &= g\left(h_m + R_m e^{j\theta}\right) \\ &= \frac{1}{\tan\left[\left(h_m + R_m e^{j\theta}\right)\frac{N}{2}\right] + \gamma\tan\left[\left(h_m + R_m e^{j\theta}\right)\frac{1}{2}\right]}, \end{aligned} \tag{23}$$

where $0 \le \theta \le 2\pi$. Hence, the m*th* root is located at

$$\omega_m = h_m + R_m\left[\frac{\int_0^{2\pi} w_m(\theta)e^{j2\theta}d\theta}{\int_0^{2\pi} w_m(\theta)e^{j\theta}d\theta}\right]. \tag{24}$$

The procedure is the same for deriving the roots of (20) with the exceptions that (23) must be modified as follows

$$w_m(\theta) = \frac{1}{\tan\left[\left(h_m + R_m e^{j\theta}\right)\frac{N}{2}\right] - \frac{1}{\gamma}\cot\left[\left(h_m + R_m e^{j\theta}\right)\frac{1}{2}\right]}, \tag{25}$$

and a suitable choice for the closed path C is a circle of radius $R_m = \pi/2N$ centered at

$$h_m = \begin{cases} (m - 1/2)\pi/N & m \le 2 \\ (m-1)\pi/N & m > 2 \end{cases}, \tag{26}$$

that can be determined by generating a plot similar to the one displayed in in Fig. 1. As an example, roots $\{\omega_k\}$ of the transcendental tangent equation, calculated by using (24) for $\rho = 0.95$ and $N = 4,\ 8,\ 16$ are tabulated in Table 1.

The implementation of the novel method to derive an explicit KLT kernel of dimension N for an arbitrary discrete data set modeled as an AR(1) is summarized as follows.

1. Estimate the first order correlation coefficient $\rho = R_{xx}(1)/R_{xx}(0) = E\{x(n)x(n+1)\}/E\{x(n)x(n)\}$ of AR(1) model for the given data set $\{x(n)\}$ where n is the index of random variables (or discrete-time) and $-1 < \rho < 1$.

2. Calculate the positive roots $\{\omega_k\}$ of the polynomial given in (17) by substituting (23) and (25) into (24) for even and odd values of k, respectively, and use the following indexing

$$m = \begin{cases} k/2 + 1 & k \text{ even} \\ (k+1)/2 & k \text{ odd} \end{cases}. \tag{27}$$

   The MATLAB™ code given in Alg. 1 shows the simplicity of this root finding method to solve a transcendental equation. It is observed from (24) and Alg. 1 (last line) that we do not need all DFT (FFT) coefficients to solve the problem since it requires only two Fourier series coefficients [8]. Therefore, it is possible to further improve the computational cost of the root finding method displayed in Alg. 1 by employing a discrete summation operator. Hence, it will have a computational complexity of $O(N)$ instead of $O(NlogN)$ required for FFT algorithms.

3. Plug in the values of ρ and $\{\omega_k\}$ in (15) and (18) to calculate the eigenvalues λ_k and eigenvectors defining the KLT matrix $\mathbf{A}_{KLT}$, respectively.

Algorithm 1 MATLAB^{TM} code of the method to calculate roots of transcendental equation given in (20). For $\rho = 0.95$, first root is calculated as 0.109447778298128 using this function. Corresponding eigenvalue calculated from (15) is 7.030310314016490. The numerical DQ algorithm [9] calculates it as 7.030310314016507.

```
rho = 0.95; % Corr.  Coeff.
N = 8; % Transform Size
L = 1024; % FFT Size
m = 1; % Root Index
if m <= 2; h=pi/N*(m-1/2);
else h = pi/N*(m-1);end
R = pi/N/2;
t = linspace(0, 2*pi*(1-1/L), L); % Theta
z = h+R*exp(1i*t); % 1i is the imaginary unit
gamma = (1 + rho)/(1-rho);
w = 1./(tan(z*N)-1/gamma*cot(z));
W = fft(conj(w),L);
omega = 2*(h + R*W(3)/W(2));
```

Remark 1: The computational cost of the proposed method to derive KLT matrix of size $N \times N$ for an arbitrary signal source has two distinct components. Namely, the calculation of the estimated first order correlation coefficient ρ for the given signal set, and the calculation of the roots $\{\omega_k\}$ of (17) that are plugged in (18) to generate the resulting transform matrix $\mathbf{A}_{KLT}$.

Remark 2: Other processes like higher order AR, auto regressive moving average (ARMA), and moving average (MA) can also be approximated by using AR(1) modeling [11]. Therefore, the proposed method to drive explicit KLT kernel may also be beneficial for other random processes of interest.

Remark 3: It was reported that the signal independent DCT kernel is identical to the KLT kernel of discrete AR(1) process in the limit when $\rho \to 1$ [12].

5. KERNEL DERIVATION EFFICIENCY

Now, we compare the computational cost of generating KLT kernel for the given statistics by employing a widely used numerical algorithm called divide and conquer (DQ) [9] and the forwarded explicit method expressed in (18). In addition, we measure the discrepancy between the kernels generated by the two competing derivation methods. A distance metric between the two kernels is defined as follows

$$d_N = \left\| \mathbf{A}^{*T}_{KLT,DQ} \mathbf{A}_{KLT,DQ} - \mathbf{A}^{*T}_{KLT,E} \mathbf{A}_{KLT,E} \right\|_2, \quad (28)$$

where $\|\cdot\|_2$ is the 2-norm, $\mathbf{A}_{KLT,DQ}$ and $\mathbf{A}_{KLT,E}$ are $N \times N$ KLT kernels obtained by using DQ and the new explicit derivation method (18), respectively. Note that the performance of the proposed method in terms of precision and derivation speed highly depends on the FFT size used in evaluating (24). Therefore, the distance metric, d_N, of (28) and the time it takes to calculate the kernel by using (18) are affected by the FFT size. Computation times (in seconds) to generate $\mathbf{A}_{KLT,DQ}$ and $\mathbf{A}_{KLT,E}$ (FFT sizes of $L = 256, 512, 1024$) for the case of $\rho = 0.95$ and $16 \leqslant N \leqslant 1024$ are displayed in Fig. 2.a. Both computations are performed by using one thread on a single processor. It is observed from Fig. 2.a that the proposed method significantly outperforms the DQ algorithm for larger values of N. Moreover, corresponding distances, d_N, measured with (28) for different N and FFT sizes are displayed in Fig. 2.b. They show that the proposed method is much faster than the currently used numerical methods with negligible discrepancy between the two kernels.

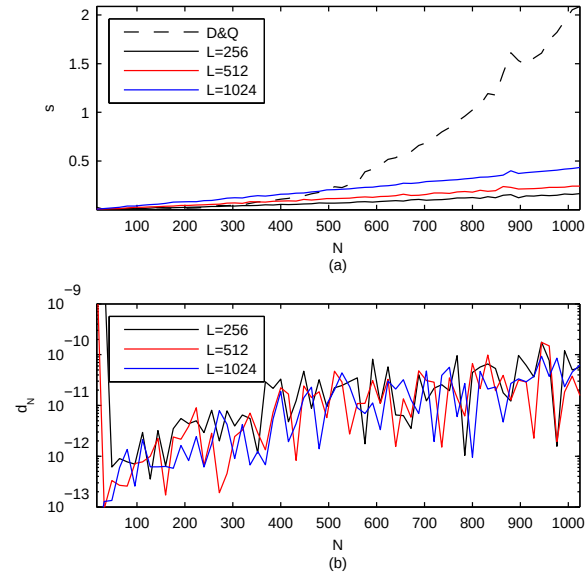

Fig. 2. (a) Computation time, in seconds, to calculate $\mathbf{A}_{KLT,DQ}$ and $\mathbf{A}_{KLT,E}$ (with FFT sizes, $L = 256, 512, 1024$) for $\rho = 0.95$ and $16 \leqslant N \leq 1024$, (b) Corresponding discrepancies, d_N, measured by using (28) for different N and FFT sizes.

6. CONCLUSIONS

A simple method to derive explicit KLT kernel for discrete AR(1) process is introduced in this paper. Its analytical framework and implementation details are presented. The merit of the proposed technique is highlighted by performance comparison with the widely used numerical DQ algorithm. The proposed method can be easily implemented on devices with highly parallel computing architectures such as field programmable gate array (FPGA) and graphics processing unit (GPU) with further performance improvements, in particular, for real-time and data intensive applications.

7. REFERENCES

[1] A. N. Akansu and R. A. Haddad, *Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets.* Academic Press, Inc., 1992.

[2] J. Wilkinson, *The Algebraic Eigenvalue Problem*. Oxford University Press, 1965.

[3] W. Ray and R. Driver, "Further decomposition of the Karhunen-Loève series representation of a stationary random process," *IEEE Transactions on Information Theory*, vol. 16, pp. 663 – 668, Nov. 1970.

[4] V. Pugachev, "A method for the determination of the eigenvalues and eigenfunctions of a certain class of linear integral equations," *Journal of Applied Mathematics and Mechanics (Translation of the Russian Journal Prikladnaya Matematika i Mekhanika)*, vol. 23, no. 3, pp. 527–533, 1959.

[5] V. Pugachev, "A method of solving the basic integral equation of statistical theory of optimum systems in finite form," *Journal of Applied Mathematics and Mechanics (Translation of the Russian Journal Prikladnaya Matematika i Mekhanika)*, vol. 23, no. 1, pp. 3–14, 1959.

[6] W. B. Davenport and W. L. Root, *An Introduction to the Theory of Random Signals and Noise*. McGraw-Hill, 1958.

[7] R. Luck and J. Stevens, "Explicit solutions for transcendental equations," *SIAM Review*, vol. 44, pp. 227–233, 2002.

[8] M. U. Torun and A. N. Akansu, "An efficient method to derive explicit KLT kernel for first-order autoregressive discrete process," *IEEE Transactions on Signal Processing, submitted*, 2012.

[9] G. Golub and C. Loan, *Matrix Computations*. Johns Hopkins University Press, 1996.

[10] M. Allen and E. Isaacson, *Numerical Analysis for Applied Science*. Pure and Applied Mathematics, John Wiley & Sons, 1997.

[11] S. Kay, *Modern Spectral Estimation: Theory and Application*. Prentice-Hall, 1988.

[12] R. J. Clarke, "Relation between the Karhunen Loève and cosine transforms," *IEE Proceedings F*, vol. 128, pp. 359 – 360, Nov. 1981.