Low Rank Circulant Approximation

by

Moody T. Chu

joined with

Robert J. Plemmons

September 12, 2000

Outline

- Background
  - ◇ Representing a Circulant Matrix
  - ◇ Basic Properties
  - ◇ Spectral Properties
  - ◇ (Inverse) Eigenvalue Problem
  - ◇ Conjugate Even Property
- Low Rank Approximation
  - ◇ TSVD
  - ◇ Data Matching Problem
  - ◇ Tree Representation
  - ◇ New Truncation Criteria
- Numerical Experiment
  - ◇ Reorganizing Tree Topology
  - ◇ Counter-intuitive TSVD
- Conclusion

Structured Low Rank Approximation

- Given
  - ◇ A target matrix $A \in R^{n\times n}$,
  - ◇ An integer k, $1 \le k < \text{rank}(A)$,
  - ◇ A class of matrices Ω with linear structure,
  - ◇ a fixed matrix norm $\|\cdot\|$;

  Find
  - ◇ A matrix $\hat{B} \in \Omega$ of rank k, and
  - ◇ $\|A - \hat{B}\| = \min_{B\in\Omega, \text{rank}(B)=k} \|A - B\|$.
- Example of linear structure:
  - ◇ Toeplitz or block Toeplitz matrices.
  - ◇ Hankel or banded matrices.
  - ◇ Circulant matrices.
- Applications:
  - ◇ Signal and image processing with Toeplitz structure.
  - ◇ Model reduction problem in speech encoding and filter design with Hankel structure.
  - ◇ Regularization of ill-posed inverse problems.

Representing a Circulant Matrix

- Basic form:

$$C = \begin{bmatrix} c_0 & c_1 & & \dots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \dots & c_{n-2} \\ c_{n-2} & c_{n-1} & c_0 & \dots & c_{n-3} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ c_1 & c_2 & & c_{n-1} & c_0 \end{bmatrix}$$

  - ◇ Uniquely determined by the first row c.
  - ◇ Denoted by $Circul(c)$.
  - ◇ Mainly interested in $c \in R^n$.

- Polynomial form:
  - ◇ Define

$$\Pi := \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & & 1 \\ 1 & 0 & & \dots & 0 \end{bmatrix}. \tag{1}$$

  - ◇ If $c := [c_0, \dots, c_{n-1}]$, then

$$Circul(c) = \sum_{k=0}^{n-1} c_k \Pi^k. \tag{2}$$

Basic Properties

- Rewrite

$$Circul(c) = P_c(\Pi) \tag{3}$$

  ◇ Characteristic polynomial

$$P_c(x) = \sum_{k=0}^{n-1} c_k x^k. \tag{4}$$

- Algebraic properties:
  - ◇ Closed under multiplication.
  - ◇ Commute under multiplication.
- Spectral properties:
  - ◇ Closely related to the Fourier analysis.
  - ◇ Explicit solution for the eigenvalue and the inverse eigenvalue problems.
  - ◇ FFT calculation.

More Spectral Properties

- Define
$$\Omega := \mathrm{diag}(1, \omega, \omega^2, \dots, \omega^{n-1}). \tag{5}$$
  - ◇ $\omega := \exp(\frac{2\pi i}{n})$.
- Define the Fourier matrix F where
$$F^* := \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2n-2} \\ \vdots & & & & \vdots \\ 1 & \omega^{n-1} & \omega^{n-2} & \dots & \omega \end{bmatrix}. \tag{6}$$
  - ◇ F is unitary.
- The forward shift matrix Π is unitarily diagonalizable.
$$\Pi = F^* \Omega F. \tag{7}$$
- The circulant matrix $Circul(c)$ with any given row vector c has a spectral decomposition
$$Circul(c) = F^* P_c(\Omega) F. \tag{8}$$

(Inverse) Eigenvalue Problem

- Forward problem:
  - ◇ Eigenvalues of $Circul(c)$:
$$\lambda = [P_c(1), \dots P_c(\omega^{n-1})]. \tag{9}$$
  - ◇ Can be computed from
$$\lambda^T = \sqrt{n} F^* c^T. \tag{10}$$
- Inverse problem:
  - ◇ Given any vector $\lambda := [\lambda_0, \dots, \lambda_{n-1}] \in C^n$, define
$$c^T = \frac{1}{\sqrt{n}} F \lambda^T. \tag{11}$$
  - ◇ $Circul(c)$ has eigenvalue λ.
- Both matrix-vector multiplication involved can be done via the fast Fourier transform (FFT).
  - ◇ Overhead is $O(n \log_2 n)$ flops.
- If all the eigenvalues are distinct, then there are precisely $n!$ many distinct circulant matrices with the prescribed spectrum.

Real Circulant Matrix

- $c^T = \frac{1}{\sqrt{n}} F \lambda^T$ is real if and only if $\lambda^T = \sqrt{n} F^* c^T$ is conjugate even.
  - ◇ If $n = 2m$, $\lambda = [\lambda_0, \lambda_1, \ldots, \lambda_{m-1}, \lambda_m, \overline{\lambda_{m-1}}, \ldots, \overline{\lambda_1}]$.
    - ▷ $\lambda_0, \lambda_m \in R$. (Absolutely real.)
  - ◇ If $n = 2m + 1$, $\lambda := [\lambda_0, \lambda_1, \ldots, \lambda_m, \overline{\lambda_m}, \ldots, \overline{\lambda_1}]$.
    - ▷ $\lambda_0 \in R$. (Absolutely real.)
- Singular value decomposition of $Circul(c)$:

$$Circul(c) = (F^* P_c(\Omega) |P_c(\Omega)|^{-1}) |P_c(\Omega)| F \tag{12}$$

  - ◇ Singular values are $|P_c(\omega^k)|$.
  - ◇ At most $\lceil \frac{n+1}{2} \rceil$ distinct singular values.

Low Rank Approximation

- Given $A \in R^{n\times n}$, its nearest circulant matrix approximation $Circul(c)$ is given by the projection

$$c_k := \frac{1}{n}\langle A, \Pi^k\rangle, \quad k = 0, \ldots, n-1, \tag{13}$$

  - ◇ $Circul(c)$ is generally of full rank even if A has lower rank to begin with.

- How to reduce the rank?
  - ◇ The truncated singular value decomposition (TSVD) gives rise to the nearest low rank approximation in Frobenius norm.
  - ◇ The TSVD of $Circul(\hat{c})$ is automatically circulant.

A Numerical Algorithm?

- Given A and rank $\ell \leq n$,
  1. Use the projection to find the nearest circulant matrix approximation $Circul(c)$ of A.
  2. Use the inverse FFT to calculate the spectrum λ of the matrix $Circul(c)$.
  3. Arrange all elements of $|\lambda|$ in descending order, including those with equal modulus.
  4. Let $\hat{\lambda}$ be the vector consisting of elements of λ, but those corresponding to the last $n - \ell$ singular values in the descending order are set to zero.
  5. Apply the FFT to $\hat{\lambda}$ to compute a nearest circulant matrix $Circul(\hat{c})$ of rank ℓ to A.
- The resulting matrix $Circul(\hat{\lambda})$ is complex-valued in general.
  - ◇ Need to preserve the conjugate even structure.
  - ◇ Need to modify the TSVD strategy.

Data Matching Problem

- All circulant matrices of the same size have the same set of unitary eigenvectors.
- The low rank real circulant approximation problem is equivalent to

  (DMP) *Given a conjugate-even vector* $\lambda \in C^n$*, find its nearest conjugate-even approximation* $\hat{\lambda} \in C^n$ *subject to the constraint that* $\hat{\lambda}$ *has exactly* $n - \ell$ *zeros.*

- How to solve DMP?
  - ⋄ Write $\hat{\lambda} = [\hat{\lambda}_1, 0] \in C^n$ with $\hat{\lambda}_1 \in C^\ell$ being arbitrary.
  - ⋄ Consider the problem of minimizing

    $$F(P, \hat{\lambda}) = \|P\hat{\lambda}^T - \lambda^T\|^2$$

    with a permutation matrix P.
    - ▷ P is used to search the match.
  - ⋄ Write $P = [P_1, P_2]$ with $P_1 \in R^{n\times\ell}$.
  - ⋄ A least squares problem:

    $$F(P, \hat{\lambda}) = \|P_1\hat{\lambda}_1^T - \lambda^T\|^2$$

◇ The optimal solution is

$$\hat{\lambda}_1 = \lambda P_1.$$

▷ The entries of $\hat{\lambda}_1$ must be a portion of λ.

◇ The objective function becomes

$$F(P, \hat{\lambda}) = \|(P_1 P_1^T - I)\lambda\|^2.$$

▷ $P_1 P_1^T - I$ is but a projection.

▷ The optimal permutation P should be such that $P_1 P_1^T$ projects λ to its first ℓ components with largest modulus.

- Without the conjugate-even constraints, the answer to the data matching problem corresponds precisely to the usual TSVD selection criterion.
- With the conjugate-even constraint, the above criterion remains effective subject to the conjugate-even structure inside λ.

An Example

- Consider the case $n = 6$.
- Assume $\lambda_1, \lambda_2 \notin$.
- Six possible conjugate-even structures.
- Tree graph:
  - ◇ Each node in the tree represents an element of λ.
  - ◇ Arrange the nodes from top to bottom in descending order of their moduli.
  - ◇ In case of a tie,
    - ▷ Complex conjugate nodes stay at the same level.
    - ▷ Real node is below the complex nodes.
- If $\lambda_1, \overline{\lambda_1}, \lambda_0, \lambda_2, \overline{\lambda_2}, \lambda_3$, then

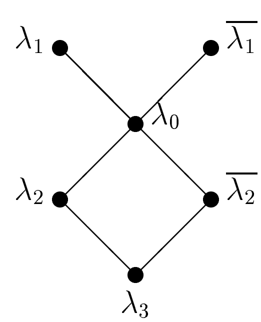

Figure 1: Tree graph of $\lambda_1, \overline{\lambda_1}, \lambda_0, \lambda_2, \overline{\lambda_2}, \lambda_3$ with $|\lambda_1| \geq |\lambda_0| > |\lambda_2| \geq |\lambda_3|$.

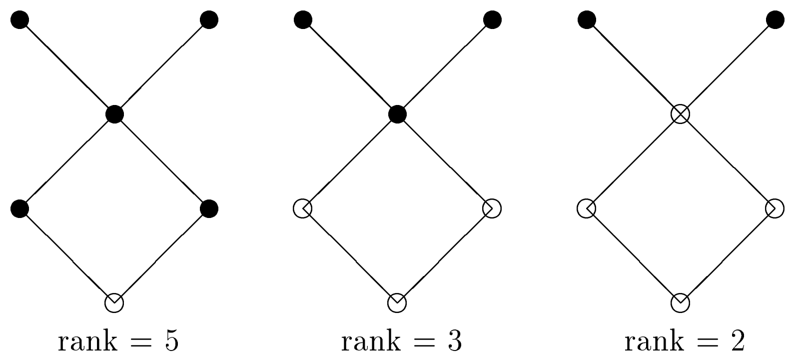

Figure 2: Tree graphs of $\hat{\lambda}$ with rank 5, 3, and 2.

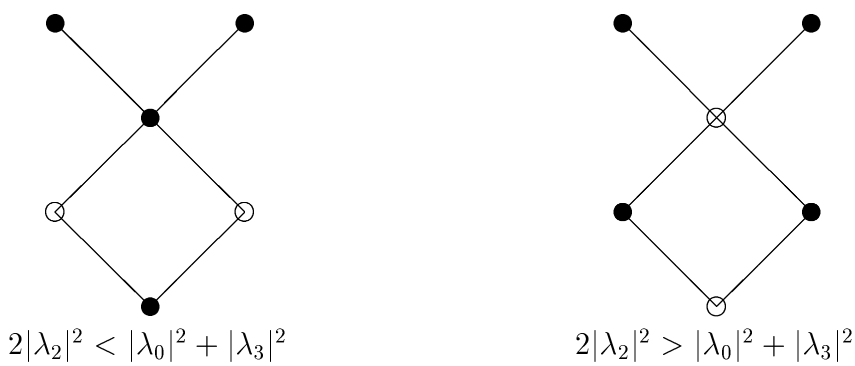

Figure 3: Tree graphs of $\hat{\lambda}$ with rank 4.

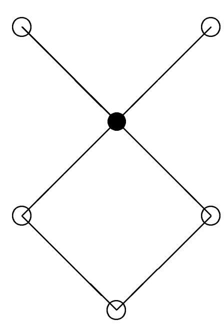

Figure 4: Tree graph of $\hat{\lambda}$ with rank 1.

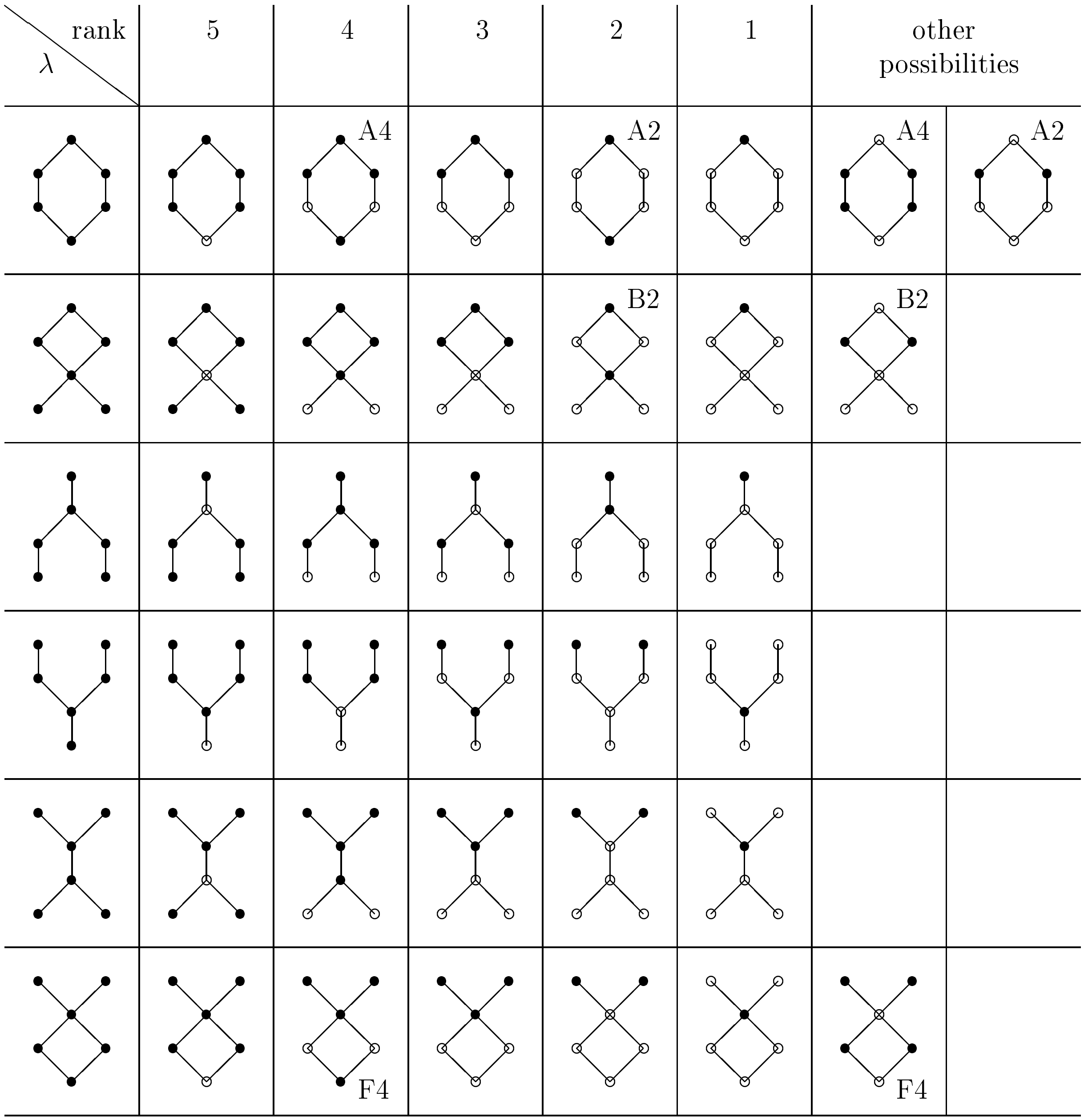

Figure 5: Possible solutions to the DMP when $n = 6$.

One More Catch

- There could be real-valued elements other than the two (when n is even) absolutely real elements in a conjugate-even λ.
  - ◇ The eigenvalues of a symmetric circulant matrix are conjugate-even and all real.
  - ◇ Non-absolutely-real, conjugate-even, real-valued elements must appear in pair.
    - ▷ The truncation criteria are further complicated.
    - ▷ The topology of the trees could be changed.
- Consider the case $n = 6$ and $\lambda_2 = \overline{\lambda_2}$. we illustrate our point below.

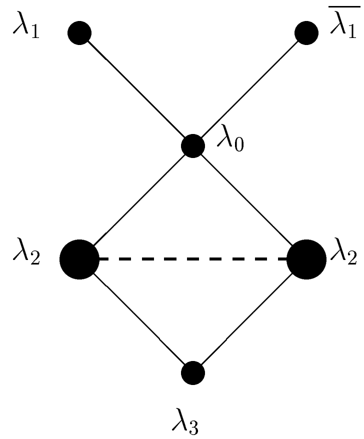

Figure 6: Tree graph of $\lambda_1, \overline{\lambda_1}, \lambda_0, \lambda_2, \lambda_2, \lambda_3$ with $|\lambda_1| \geq |\lambda_0| > |\lambda_2| \geq |\lambda_3|$.

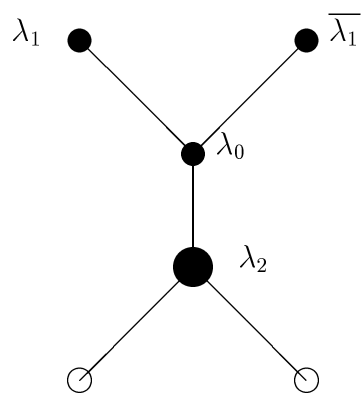

Figure 7: Tree graph of $\hat{\lambda}$ with rank 4 when $\lambda_2 = \overline{\lambda_2}$.

A Numerical Algorithm!

- For the case $n = 2m$, we have assumed
  - ◇ 2 absolutely real elements $|\lambda_0| \geq |\lambda_m|$.
  - ◇ $2m - 2$ elements are "potentially" complex-valued , that they are paired up (necessarily), and are arranged in descending order, i.e., $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_{m-1}|$.
- No ordering relationship between the absolutely real elements and the potentially complex elements is assumed.
  - ◇ Such an ordering relationship determines the truncation criteria.
  - ◇ Assuming that there are exactly $m+1$ distinct absolute values of elements in λ, then there are exactly $\binom{m+1}{2}$ many possible conjugate-even structures for the case $n = 2m$.
- Any algorithm needs to be smart enough to explore the conjugate even structure, to truncate, and to reassemble the conjugate even structure.

Example 1

Consider the 8×8 symmetric $Circul(c)$:

$$c = \big[0.5404, 0.2794, 0.1801, \text{-}0.0253, \text{-}0.2178, \text{-}0.0253, 0.1801, 0.2794\big].$$

- Eigenvalues (in descending order):

  $$\big[\mathbf{1.1909}, 1.1892, 1.1892, 0.3273, 0.3273, \mathbf{0.1746}, \text{-}0.0376, \text{-}0.0376\big]$$

- For rank 7 approximation, the usual TSVD would set -0.0376 to zero, resulting in a complex matrix.
- Use the conjugate-even eigenvalues

  $$\hat{\lambda} = \big[1.1909, 1.1892, 0.3273, \text{-}0.0376, \mathbf{0}\ \text{-}0.0376, 0.3273, 1.1892\big],$$

  to obtain the best real-valued, rank 7, approximation $Circul(\hat{c})$ via the FFT:

  $$\hat{c} = \big[0.5186, 0.3657, 0.0670, \text{-}0.0680, \text{-}0.0572, \text{-}0.0680, 0.0670, 0.3657\big].$$

- To obtain the best real-value, rank 4, circulant approximation, use eigenvalues $\hat{\lambda}$

  $$\hat{\lambda} = \big[1.1909, 1.1892, 0, 0, 0.3273, 0, 0, 1.1892\big].$$

  - ◇ The last pair of eigenvalues in λ are set to zero while the value 0.1746 together with one 0.3273 cause a topology change in the graph tree.

Example 2

Consider the 9×9 $Circul(c)$ with

$$c = \big[1.6864,\ 1.7775,\ 1.9324,\ 2.9399,\ 1.9871,\ 1.7367,\ 4.0563,\ 1.2848,\ 2.5989\big].$$

- Eigenvalues: structure given by

  [**20.0000**,

  -2.8130 + 1.9106i, -2.8130 - 1.9106i, 3.0239 - 1.0554i, 3.0239 + 1.0554i,

  -1.3997 + 0.7715i, -1.3997 - 0.7715i, -1.2223 - 0.2185i, -1.2223 + 0.2185i].

- To obtain a real-valued, rank 8, circulant approximation of rank 8, we have no choice but to select the set the *largest* eigenvalue (singular value) of $Circul(c)$ to zero to produce

  $$\hat{c} = \big[-0.5358,\ -0.5872,\ -1.1736,\ -0.3212,\ 1.0198,\ 1.4013,\ -0.0761,\ -0.4115,\ 0.6844\big].$$

  as its first row.

  - ◇ Setting the largest singular value to zero to obtain the nearest low rank approximation is quite counter-intuitive to the usual sense of TSVD.

Conclusion

- For any given real data matrix, its nearest real circulant approximation can simply be determined from the average of its diagonal entries.
- The nearest low rank approximation to a circulant matrix can be determined effectively from the TSVD and the FFT.
- To construct real circulant matrix with specified spectrum, the eigenvalues must appear in conjugate even form.
- The truncation criteria for a nearest low rank, real, circulant matrix approximation must be modified.
- We have proposed a fast algorithm to accomplish all of these objectives.
- Extensions to the block case with possible applications to image reconstruction (not discussed in this talk) are possible.