THE CIRCULANT OPERATOR IN THE BANACH ALGEBRA OF MATRICES

Raymond H. Chan, Xiao–Qing Jin and Man–Chung Yeung

Department of Mathematics

University of Hong Kong

Hong Kong

May 1990

Revised July 1990

Abstract. We study an operator c which maps every n-by-n matrix A_n to a circulant matrix $c(A_n)$ that minimizes the Frobenius norm $\|A_n - C_n\|_F$ over all n-by-n circulant matrices C_n. The circulant matrix $c(A_n)$, called the optimal circulant preconditioner, has proved to be a good preconditioner for a general class of Toeplitz systems. In this paper, we give different formulations of the operator, discuss its algebraic and geometric properties and compute its operator norms in different Banach algebras of matrices. Using these results, we are able to give an efficient algorithm for finding the super-optimal circulant preconditioner which is defined to be the minimizer of $\|I - C_n^{-1}A_n\|_F$ over all nonsingular circulant matrices C_n.

Abbreviated Title. Circulant Operator

Key Words. Toeplitz matrix, circulant matrix, optimal preconditioner, circulant operator, preconditioned conjugate gradient method

AMS(MOS) Subject Classifications. 65F10, 65F15

§1 Introduction.

Preconditioned conjugate gradient methods have been used successfully in solving many large matrix problems. Strang [6] first proposed using the method with circulant preconditioners for solving Toeplitz systems. R. Chan and Strang [1] then proved that for Toeplitz systems with generating functions that are positive functions in the Wiener class, the method has a super-linear convergence rate due to the clustering of the eigenvalues of the preconditioned matrices.

Several circulant preconditioners have been proposed since then, see for example T. Chan [4] and Tyrtyshinkov [7]. For any n-by-n matrix A_n, the circulant preconditioner proposed in T. Chan [4], called the *optimal circulant preconditioner*, is defined to be the minimizer of $\|C_n - A_n\|_F$ over the space of all n-by-n circulant matrices C_n. Here $\|\cdot\|_F$ denotes the Frobenius norm. The circulant preconditioner given in Tyrtyshinkov [7] is defined to be the minimizer of $\|I - C_n^{-1}A_n\|_F$ over the space of all nonsingular circulant matrices C_n, and is called the *super-optimal circulant preconditioner*. From the computational point of view, these optimal circulant preconditoners are better than the one proposed in Strang [6] because they are symmetric positive definite whenever A_n is. Numerical results in these papers showed that they are very good preconditioners. The analysis of the convergence rates of these preconditioned systems are given in R. Chan [2] and R. Chan et al. [3], and it is proved that for the same class of Toeplitz systems mentioned above, these methods converge at the same rate as the Strang's preconditioned systems.

In this paper, we study these circulant preconditioners from the operator point of view. Let $(\mathcal{M}_{n\times n}, \|\cdot\|)$ be the Banach algebra of all n-by-n matrices over the complex field, equipped with a matrix norm $\|\cdot\|$. Let $(\mathcal{C}_{n\times n}, \|\cdot\|)$ be the subalgebra of all circulant matrices. We note that $\mathcal{C}_{n\times n}$ is an inverse-closed, commutative algebra. Let c be an operator defined on $(\mathcal{M}_{n\times n}, \|\cdot\|)$ such that for any A_n in $\mathcal{M}_{n\times n}$, $c(A_n)$ is the minimizer of $\|A_n - C_n\|_F$ over all C_n in $\mathcal{C}_{n\times n}$. Obviously, $c(A_n)$ is the optimal circulant preconditioner proposed in T. Chan [4] and c is an operator from $(\mathcal{M}_{n\times n}, \|\cdot\|)$ into the subalgebra $(\mathcal{C}_{n\times n}, \|\cdot\|)$. We call c the *circulant operator*. In R. Chan et al. [3], we utilized this operator to analyze the convergence rate of Toeplitz systems preconditioned by super-optimal circulant preconditioners.

The purpose of this paper is to discuss some other aspects of this operator. The outline of the paper is as follows. In §2, we introduce other formulations of the operator and prove some of its algebraic and geometric properties. In §3, we compute its operator norms for different Banach algebras of matrices. In §4, we apply these results to derive an algorithm for finding the super-optimal circulant preconditioner. Our algorithm is more efficient than the one proposed in Trytyshinkov [7].

§2 The Circulant Operator.

In this section, we discuss some properties of the circulant operator. For any A_n in $\mathcal{M}_{n\times n}$, let $\delta(A_n)$ denote the diagonal matrix whose diagonal is equal to the diagonal of the matrix A_n. We first give two methods for finding $c(A_n)$.

Theorem 1. *Let $A_n = (a_{ij}) \in \mathcal{M}_{n\times n}$ and $c(A_n)$ be the minimizer of $\|C_n - A_n\|_F$ over all $C_n \in \mathcal{C}_{n\times n}$. Then $c(A_n)$ is uniquely determined by A_n. Moreover,*

(i) $c(A_n)$ is given by

$$c(A_n) = \sum_{j=0}^{n-1} \Big(\frac{1}{n} \sum_{p-q\equiv j \pmod n} a_{pq}\Big) Q^j \ , \tag{1}$$

where Q is the n-by-n circulant matrix

$$Q = \begin{bmatrix} 0 & & & & 1 \\ 1 & 0 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ \mathbf{0} & & & 1 & 0 \end{bmatrix}; \tag{2}$$

(ii) $c(A_n)$ is also given by

$$c(A_n) = F^* \delta(F A_n F^*) F, \tag{3}$$

where F is the Fourier matrix and "$*$" denotes conjugate transposition.

Proof. For the proof of (i), see Theorem 2.1 in Tyrtyshnikov [7]. For (ii), we first note that any circulant matrix C_n can be expressed as $F^*\Lambda_n F$, where Λ_n is a diagonal matrix containing the eigenvalues of C_n, see Davis [5]. Since the Frobenius norm is unitary-invariant, we have

$$\|C_n - A_n\|_F = \|F^*\Lambda_n F - A_n\|_F = \|\Lambda_n - FA_nF^*\|_F.$$

Thus the problem of minimizing $\|C_n - A_n\|_F$ over $\mathcal{C}_{n\times n}$ is equivalent to the problem of minimizing $\|\Lambda_n - FA_nF^*\|_F$ over all diagonal matrices. Since Λ_n can only affect the diagonal entries of FA_nF^*, we see that the solution for the latter problem is $\Lambda_n = \delta(FA_nF^*)$. Hence $F^*\delta(FA_nF^*)F$ is the minimizer of $\|C_n - A_n\|_F$. It is clear from the argument above that Λ_n and hence $c(A_n)$ are uniquely determined by A_n. □

We remark that by (1), the j-th entry in the first column of $c(A_n)$ is given by

$$[c(A_n)]_{j0} = \frac{1}{n}\sum_{p-q\equiv j \pmod n} a_{pq} = \frac{1}{n}\,\mathrm{tr}\,(A_nQ^{-j}), \quad j = 0, 1, \cdots, n-1, \tag{4}$$

where $\mathrm{tr}\,(\cdot)$ denotes the trace. By (3), the eigenvalues of $c(A_n)$ are given by $\delta(FA_nF^*)$. We also notice that the nonsingularity of A_n cannot guarantee $\delta(FA_nF^*)$ to be nonsingular. Hence $c(A_n)$ may be singular for nonsingular A_n.

The following Lemma is on the algebraic properties of the circulant operator.

Lemma 1.

(i) For all A_n, $B_n \in \mathcal{M}_{n\times n}$ and α, β complex scalars, $c(\alpha A_n + \beta B_n) = \alpha c(A_n) + \beta c(B_n)$. Moreover, for all $A_n \in \mathcal{M}_{n\times n}$, $c^2(A_n) = c(c(A)) = c(A_n)$. Thus c is a linear projection operator.

(ii) Let $A_n \in \mathcal{M}_{n\times n}$, tr $(c(A_n)) =$ tr $(A_n) = \sum_{j=0}^{n-1} \lambda_j(A_n)$, where $\lambda_j(A_n)$ are the eigenvalues of A_n.

(iii) *For all* $A_n \in \mathcal{M}_{n\times n}$*, we have* $c(A_n^*) = c(A_n)^*$.

(iv) *Let* $A_n \in \mathcal{M}_{n\times n}$ *and* $C_n \in C_{n\times n}$*. Then*

$$c(C_n A_n) = C_n \cdot c(A_n),$$

$$c(A_n C_n) = c(A_n) \cdot C_n.$$

Proof. The proofs of (i) and (ii) are trivial, therefore we omit them. By using (3) and the fact that $\delta(A_n^*) = (\delta(A_n))^*$, one can easily prove (iii). For the proof of (iv), see Theorem 2 in R. Chan et al. [3]. □

Next we are going to give some geometric properties of the circulant operator. For all A_n , $B_n \in \mathcal{M}_{n\times n}$, let $\langle A_n, B_n\rangle_F \equiv \frac{1}{n}$ tr $(A_n B_n^*)$. Obviously $\langle A_n, B_n\rangle_F$ is an inner product in $\mathcal{M}_{n\times n}$ and $\langle A_n, A_n\rangle_F = \frac{1}{n}\|A_n\|_F^2$. It is easy to show that $\{Q^j \mid j = 0, \cdots, n-1\}$, where Q is given in (2), is an orthonormal basis of $(\mathcal{C}_{n\times n}, \|\cdot\|_F)$. We show below that $A_n - c(A_n)$ is perpendicular to the space $\mathcal{C}_{n\times n}$.

Lemma 2. *Let* $A_n \in \mathcal{M}_{n\times n}$*, then we have*

(i) $\langle A_n - c(A_n), C_n\rangle_F = 0$ *for all* $C_n \in \mathcal{C}_{n\times n}$,

(ii) $\langle A_n, c(A_n)\rangle_F = \frac{1}{n}\|c(A_n)\|_F^2$,

(iii) $\|A_n - c(A_n)\|_F^2 = \|A_n\|_F^2 - \|c(A_n)\|_F^2$.

Proof. For (i), since $\{Q^j\}_{j=0}^{n-1}$ is an orthonormal basis of $\mathcal{C}_{n\times n}$, it suffices to show that $\langle A_n - c(A_n), Q^j\rangle_F = 0$ for $j = 0, \cdots, n-1$. However, by (4) and Lemma 1 (i), we have

$$\begin{aligned}\langle A_n - c(A_n), Q^j\rangle_F &= \frac{1}{n} \text{ tr } [(A_n - c(A_n))Q^{-j}] = \frac{1}{n} \text{ tr } (A_n Q^{-j}) - \frac{1}{n} \text{ tr } (c(A_n)Q^{-j}) \\ &= [c(A_n)]_{j0} - [c(c(A_n))]_{j0} = [c(A_n)]_{j0} - [c(A_n)]_{j0} = 0.\end{aligned}$$

Now (ii) follows directly from (i). For (iii), we have, by parts (i) and (ii) above,

$$\begin{aligned}\|A_n - c(A_n)\|_F^2 &= n\langle A_n - c(A_n), A_n - c(A_n)\rangle_F = n\langle A_n - c(A_n), A_n\rangle_F \\ &= n\langle A_n, A_n\rangle_F - n\langle c(A_n), A_n\rangle_F = \|A_n\|_F^2 - \|c(A_n)\|_F^2. \quad \square\end{aligned}$$

§3 Spectral Properties of the Circulant Operator.

In this section, we discuss some spectral properties of the circulant operator. The following theorem was first proved for the real scalar field in Tyrtyshnikov [7]. His proof uses equation (1) and our proof here uses equation (3).

Theorem 2. *If A_n is Hermitian, then $c(A_n)$ is Hermitian. Moreover, we have*

$$\lambda_{\min}(A_n) \le \lambda_{\min}(c(A_n)) \le \lambda_{\max}(c(A_n)) \le \lambda_{\max}(A_n) \ ,$$

where $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ denote the largest and the smallest eigenvalues respectively. In particular, if A_n is positive definite, then $c(A_n)$ is also positive definite.

Proof. By Lemma 1 (iii), it is clear that $c(A_n)$ is Hermitian when A_n is Hermitian. By (3), we know that the eigenvalues of $c(A_n)$ are given by $\delta(FA_nF^*)$. Suppose that $\delta(FA_nF^*) = \operatorname{diag}(\lambda_0, \cdots, \lambda_{n-1})$ with $\lambda_j = \lambda_{\min}(c(A_n))$ and $\lambda_k = \lambda_{\max}(c(A_n))$. Let e_j and e_k denote the j-th and the k-th unit vectors respectively. Since A_n is Hermitian, we have

$$\lambda_{\max}(c(A_n)) = \lambda_k = \frac{e_k^* F A_n F^* e_k}{e_k^* e_k} \le \max_{x \neq 0} \frac{x^* F A_n F^* x}{x^* x} = \max_{x \neq 0} \frac{x^* A_n x}{x^* x} = \lambda_{\max}(A_n).$$

Similarly,

$$\lambda_{\min}(A_n) = \min_{x \neq 0} \frac{x^* A_n x}{x^* x} = \min_{x \neq 0} \frac{x^* F A_n F^* x}{x^* x} \le \frac{e_j^* F A_n F^* e_j}{e_j^* e_j} = \lambda_j = \lambda_{\min}(c(A_n)).$$

From the inequality above, we can easily see that $c(A_n)$ is positive definite when A_n is positive definite. □

Lemma 3. *For all $A_n \in \mathcal{M}_{n \times n}$, $c(A_n A_n^*) - c(A_n)c(A_n^*)$ is a semi-positive definite matrix.*

Proof. Let $A_n = (a_{jk})$ and $[F]_{jk} = \frac{1}{\sqrt{n}} \xi_j^k$, where $\xi_j = e^{-\frac{2\pi i j}{n}}$. If we let

$$D_n = c(A_n A_n^*) - c(A_n)c(A_n^*) = F^*(\delta(FA_nA_n^*F^*) - \delta(FA_nF^*)\delta(FA_n^*F^*))F,$$

then for all $k = 0, \cdots, n-1$, we have

$$[\delta(FA_nA_n^*F^*)]_{kk} = [\delta((FA_n)(FA_n)^*)]_{kk} = \frac{1}{n} \sum_{q=0}^{n-1} (\sum_{p=0}^{n-1} a_{pq} \xi_k^p) \overline{(\sum_{p=0}^{n-1} a_{pq} \xi_k^p)},$$

and

$$[\delta(FA_nF^*)\delta(FA_n^*F^*)]_{kk} = (\frac{1}{n}\sum_{p=0}^{n-1}\sum_{q=0}^{n-1} a_{pq}\xi_k^{p-q})\overline{(\frac{1}{n}\sum_{p=0}^{n-1}\sum_{q=0}^{n-1} a_{pq}\xi_k^{p-q})}.$$

Hence the k-th eigenvalue of D_n is given by

$$\lambda_k(D_n) = \frac{1}{n}\sum_{q=0}^{n-1}|\sum_{p=0}^{n-1} a_{pq}\xi_k^p|^2 - |\frac{1}{n}\sum_{p=0}^{n-1}\sum_{q=0}^{n-1} a_{pq}\xi_k^{p-q}|^2 .$$

Since

$$|\frac{1}{n}\sum_{p=0}^{n-1}\sum_{q=0}^{n-1} a_{pq}\xi_k^{p-q}| \le \frac{1}{n}\sum_{q=0}^{n-1}|\sum_{p=0}^{n-1} a_{pq}\xi_k^p||\xi_k^{-q}| = \frac{1}{n}\sum_{q=0}^{n-1}|\sum_{p=0}^{n-1} a_{pq}\xi_k^p|,$$

we have

$$\lambda_k(D_n) \ge \frac{1}{n}\sum_{q=0}^{n-1}|\sum_{p=0}^{n-1} a_{pq}\xi_k^p|^2 - (\frac{1}{n}\sum_{q=0}^{n-1}|\sum_{p=0}^{n-1} a_{pq}\xi_k^p|)^2 .$$

Let $d_{qk} = \frac{1}{n}|\sum_{p=0}^{n-1} a_{pq}\xi_k^p|$, then by Cauchy-Schwartz inequality, we have

$$\lambda_k(D_n) \ge n\sum_{q=0}^{n-1} d_{qk}^2 - (\sum_{q=0}^{n-1} d_{qk})^2 \ge 0, \quad k = 0, \cdots, n-1.$$

Thus D_n is semi-positive definite. □

Theorem 3. *For all $n \ge 1$, we have*

(i) $\|c\|_1 \equiv \sup_{\|A_n\|_1=1} \|c(A_n)\|_1 = 1$,

(ii) $\|c\|_\infty \equiv \sup_{\|A_n\|_\infty=1} \|c(A_n)\|_\infty = 1$,

(iii) $\|c\|_F \equiv \sup_{\|A_n\|_F=1} \|c(A_n)\|_F = 1$,

(iv) $\|c\|_2 \equiv \sup_{\|A_n\|_2=1} \|c(A_n)\|_2 = 1$.

Proof. To prove (i), we first note that if $A_n = I$, then $\|c(A_n)\|_1 = \|I\|_1 = 1$. For general A_n in $\mathcal{M}_{n\times n}$, we have by (1)

$$\begin{aligned}\|c(A_n)\|_1 &= \sum_{j=0}^{n-1}\left|\frac{1}{n}\sum_{p-q\equiv j \pmod n} a_{pq}\right| \le \frac{1}{n}\sum_{j=0}^{n-1}\sum_{p-q\equiv j \pmod n}|a_{pq}| \\ &= \frac{1}{n}\sum_{i=0}^{n-1}\sum_{k=0}^{n-1}|a_{ik}| \le \frac{1}{n}\cdot n\cdot\|A_n\|_1.\end{aligned}$$

Hence $\|c\|_1 = 1$ for all n. By a similar argument, we can prove (ii).

To prove (iii), we notice that if $A_n = \frac{1}{n}I$, then $\|c(A_n)\|_F = \frac{1}{n}\|I\|_F = 1$. For general A_n in $\mathcal{M}_{n\times n}$, by Lemma 2 (iii), we have

$$\|c(A_n)\|_F^2 = \|A_n\|_F^2 - \|A_n - c(A_n)\|_F^2 \leq \|A_n\|_F^2.$$

Thus $\|c(A_n)\|_F \leq \|A_n\|_F$. Hence $\|c\|_F = 1$ for all n.

To prove (iv), by Lemma 1 (iii), Lemma 3 and Theorem 2, we have

$$\begin{aligned}\|c(A_n)\|_2^2 &= \lambda_{\max}(c(A_n)^* c(A_n)) = \lambda_{\max}(c(A_n^*)c(A_n)) \\ &\leq \lambda_{\max}(c(A_n^* A_n)) \leq \lambda_{\max}(A_n^* A_n) = \|A_n\|_2^2,\end{aligned}$$

for all A_n in $\mathcal{M}_{n\times n}$. Since $\|c(I)\|_2 = \|I\|_2 = 1$, $\|c\|_2 = 1$. □

§4 The Super-optimal Circulant Preconditioner.

In this section, we apply the results in previous sections to analyze the super-optimal circulant preconditioner proposed in Tyrtyshnikov [7]. For A_n in $\mathcal{M}_{n\times n}$, the preconditioner is defined to be the minimizer of $\|I - C_n^{-1}A_n\|_F$ over all nonsingular $C_n \in \mathcal{C}_{n\times n}$. First, we generalize Thoerem 4.1 in Tyrtyshnikov [7] from the real field to the complex field. As with Theorem 2, his proof uses equation (1) and ours uses equation (3).

Theorem 4. *Let $A_n \in \mathcal{M}_{n\times n}$ be such that both A_n and $c(A_n)$ are nonsingular. Then the super-optimal circulant preconditioner for A_n exists and is equal to $c(A_nA_n^*)c(A_n^*)^{-1}$.*

Proof. Instead of minimizing $\|I - C_n^{-1}A_n\|_F$, we consider the problem of minimizing $\|I - \widehat{C}_nA_n\|_F$ over all nonsingular $\widehat{C}_n$ in $\mathcal{C}_{n\times n}$. Letting $\widehat{C}_n = F^*\Lambda_nF$, we have

$$\begin{aligned}\|I - \widehat{C}_nA_n\|_F =& \|I - F^*\Lambda_nFA_n\|_F = \|I - \Lambda_nFA_nF^*\|_F \\ &= \text{tr } (I - \Lambda_nFA_nF^* - FA_n^*F^*\Lambda_n^* + \Lambda_nFA_nA_n^*F^*\Lambda_n^*) \\ &= \text{tr } (I - \Lambda_n\delta(FA_nF^*) - \delta(FA_n^*F^*)\Lambda_n^* + \Lambda_n\delta(FA_nA_n^*F^*)\Lambda_n^*).\end{aligned}$$

Let Λ_n, $\delta(FA_nF^*)$ and $\delta(FA_nA_n^*F^*)$ be given by $\text{diag}(\lambda_0, \cdots, \lambda_{n-1})$, $\text{diag}(u_0, \cdots, u_{n-1})$ and $\text{diag}(w_0, \cdots, w_{n-1})$ respectively. We have

$$\begin{aligned}
\min \|I - \widehat{C}_n A_n\|_F &= \min \{ \text{ tr } [I - \Lambda_n \delta(FA_nF^*) - \delta(FA_n^*F^*)\Lambda_n^* + \Lambda_n \delta(FA_nA_n^*F^*)\Lambda_n^*] \} \\
&= \min_{\{\lambda_0, \cdots, \lambda_{n-1}\}} \sum_{k=0}^{n-1} (1 - \lambda_k u_k - \overline{u}_k \overline{\lambda}_k + \lambda_k w_k \overline{\lambda}_k).
\end{aligned}$$

Notice that by (3) and Lemma 3, $w_k \geq u_k \bar{u}_k$ for all $k = 0, \cdots, n-1$. Hence for all complex scalars λ_k, $k = 0, \cdots, n-1$, the terms $1 - \lambda_k u_k - \overline{u}_k \overline{\lambda}_k + \lambda_k w_k \overline{\lambda}_k$ are non-negative. Differentiating them with respect to the real and imaginary parts of λ_k and setting the derivatives to zero, we get

$$\lambda_k = \frac{\overline{u}_k}{w_k}, \qquad k = 0, \cdots, n-1.$$

Since A_n and $c(A_n)$ are nonsingular, both w_k and u_k are nonzero. Hence λ_k are also nonzero. Thus the minimizer of $\|I - \widehat{C}_n A_n\|_F$ is nonsingular and is given by

$$\begin{aligned}
\widehat{C}_n &= F^* \Lambda_n F = F^* \delta(FA_n^*F^*)[\delta(FA_nA_n^*F^*)]^{-1} F \\
&= (F^* \delta(FA_n^*F^*)F)(F^* \delta(FA_nA_n^*F^*)F)^{-1} = c(A_n^*) c(A_n A_n^*)^{-1}.
\end{aligned}$$

Therefore the super-optimal circulant preconditioner is given by $\widehat{C}_n^{-1} = c(A_nA_n^*)c(A_n^*)^{-1}$.

□

We remark that by Theorem 2, if A_n is Hermitian positive definite, then $c(A_n)$ is nonsingular. Hence the super-optimal circulant preconditioner is defined for all Hermitian positive definite matrices.

When the system $A_n x = b$ is solved by preconditioned conjugate gradient method with the super-optimal circulant preconditioner $c(A_nA_n^*)c(A_n^*)^{-1}$, then in each iteration, we have to compute a matrix-vector multiplication of the form $c(A_n^*)c(A_nA_n^*)^{-1}y$. We now derive an algorithm for finding $c(A_n^*)c(A_nA_n^*)^{-1}$. We begin by considering a general A_n that has no special structure. We first note that $c(A_n^*)c(A_nA_n^*)^{-1} = \widehat{C}_n$ is circulant.

Hence it is determined by its first column, which is given by

$$\begin{aligned}\widehat{C}_n e_0 &= c(A_n^*)[c(A_n A_n^*)]^{-1} e_0 = F^* \delta(F A_n^* F^*)[\delta(F A_n A_n^* F^*)]^{-1} F e_0 \\ &= F^* \delta(F A_n^* F^*)[\delta(F A_n A_n^* F^*)]^{-1} \mathbf{1}. \end{aligned} \tag{5}$$

Here $\mathbf{1}$ is the vector of all ones. To compute $\delta(FA_n^*F^*)$, it is clear from (1) that the first column $c(A_n^*)e_0$ of $c(A_n^*)$ can be computed in n^2 additions and n multiplications. Since by (3), $\delta(FA_n^*F^*)\mathbf{1} = Fc(A_n^*)e_0$, one FFT is required to obtain $\delta(FA_n^*F^*)$. To compute $\delta(FA_nA_n^*F^*) = \delta((FA_n)(FA_n)^*)$, we first need n FFTs to get FA_n, then another n^2 additions and n^2 multiplications to obtain the diagonal entries of $\delta((FA_n)(FA_n)^*)$. Now $\delta(FA_n^*F^*)[\delta(FA_nA_n^*F^*)]^{-1}$ can be obtained by n multiplications. By (5), one additional FFT is required to get $\widehat{C}_n e_0$. Thus for an arbitrary n-by-n matrix A_n, $\widehat{C}_n$ can be computed within $2n^2$ additions, $2n+n^2$ multiplications and $(n+2)$ FFTs.

We remark that from the computational point of view, we do not require the explicit form of $\widehat{C}_n$, we only need its eigenvalues and they are given by the diagonal entries of $\delta(FA_n^*F^*)[\delta(FA_nA_n^*F^*)]^{-1}$. In fact, given any vector y, $\widehat{C}_n y$ can be computed by

$$\widehat{C}_n y = F^* \delta(FA_n^*F^*)[\delta(FA_nA_n^*F^*)]^{-1} F y.$$

Hence the last FFT in the above algorithm can usually be saved.

Next we study how a Toeplitz structure can be exploited to accelerate the computation of $\widehat{C}_n$. The algorithm presented here is more efficient than the one proposed in Trytyshinov [7] where a Toeplitz matrix is partitioned into the sum of low and upper triangular Toeplitz matrices. Here we will partition a Toeplitz matrix into the sum of a circulant matrix C_n and a skew-circulant matrix S_n. Let $A = (a_{ij}) = (a_{i-j})$ be Toeplitz, define C_n and S_n by

$$C_n = \frac{1}{2}\begin{bmatrix} a_0 & a_{-1}+a_{n-1} & & & a_{-(n-1)}+a_1 \\ a_1+a_{-(n-1)} & a_0 & & & \\ & & \ddots & \ddots & \\ & & \ddots & \ddots & \\ a_{n-1}+a_{-1} & & & & a_0 \end{bmatrix},$$

and

$$S_n = \frac{1}{2}\begin{bmatrix} a_0 & a_{-1}-a_{n-1} & & & a_{-(n-1)}-a_1 \\ -(a_{-(n-1)}-a_1) & a_0 & & & \\ & & \ddots & \ddots & \\ & & \ddots & \ddots & \\ -(a_{-1}-a_{n-1}) & & & & a_0 \end{bmatrix}.$$

Clearly C_n is circulant and S_n is skew-circulant. Moreover, we have $A_n = C_n + S_n$ and that $C_n e_0$ and $S_n e_0$ can be computed by $2n$ multiplications and $2n$ additions. We remark that $2 \cdot C_n$ is the circulant preconditioner proposed in R. Chan [2].

We will compute the first column of $\widehat{C}_n$ by (5). We first compute $\delta(FA_nA_n^*F^*)$. Since $A_n = C_n + S_n = F^*\Lambda_c F + S_n$, where Λ_c is the diagonal matrix containing the eigenvalues of C_n, we have

$$FA_nF^* = \Lambda_c + FS_nF^*. \tag{6}$$

Hence

$$\begin{aligned}\delta(FA_nA_n^*F^*) =& \delta((FA_nF^*)(FA_nF^*)^*) = \delta((\Lambda_c + FS_nF^*)(\Lambda_c^* + FS_n^*F^*)) \\ =& \Lambda_c\Lambda_c^* + \delta(FS_nF^*)\Lambda_c^* + \Lambda_c\delta(FS_n^*F^*) + \delta(FS_nS_n^*F^*). \end{aligned} \tag{7}$$

We now consider the terms in the right hand side of (7) one by one.

(i) For the first term in (7), we first compute Λ_c by using $\Lambda_c\mathbf{1} = \Lambda_c Fe_0 = FC_ne_0$. That requires one FFT. Then $\Lambda_c\Lambda_c^*$ can be computed in n multiplications.

(ii) For $\delta(FS_nF^*)\Lambda_c^*$, we know that by (3),

$$\delta(FS_nF^*)\mathbf{1} = \delta(FS_nF^*)Fe_0 = Fc(S_n)e_0.$$

Since S_n is skew-circulant, $c(S_n)e_0$ can be computed in $3n$ multiplications and n additions. Then $\delta(FS_nF^*)\Lambda_c^*$ can be obtained by an additional n multiplications and one FFT.

(iii) The third term in (7) is just the conjugate transpose of the second term in (7). Hence it can be computed without any work at all.

(iv) Finally for $\delta(FS_nS_n^*F^*)$, we have by (3) again,

$$\delta(FS_nS_n^*F^*)\mathbf{1} = \delta(FS_nS_n^*F^*)Fe_0 = Fc(S_nS_n^*)e_0. \tag{8}$$

Thus the main work is to compute $c(S_nS_n^*)e_0$. We first find $S_nS_n^*$. We note that for all skew-circulant matrices, and in particular for S_n, they can be written as

$$S_n = \Theta^* F^* \Lambda_s F \Theta, \tag{9}$$

where $\Theta = \operatorname{diag}(1, e^{\frac{\pi}{n}i}, \cdots, e^{\frac{(n-1)\pi}{n}i})$ and Λ_s is the diagonal matrix containing the eigenvalues of S_n, see for instance, Davis [5]. Because $\Lambda_s\mathbf{1} = \Lambda_s F\Theta e_0 = F\Theta S_n e_0$, Λ_s can be computed in n multiplications and one FFT. Since $S_nS_n^*$ is still skew-circulant, it is determined by its first column $S_nS_n^*e_0$. By (9),

$$S_nS_n^*e_0 = \Theta^*F^*\Lambda_s\Lambda_s^*F\Theta e_0 = \Theta^*F^*\Lambda_s\Lambda_s^*\mathbf{1},$$

which can be computed by using one FFT and $2n$ multiplications. Once we know $S_nS_n^*e_0$, $c(S_nS_n^*)e_0$ can be computed by using another $3n$ multiplications and n additions. Finally by (8), one additional FFT is required to get $\delta(FS_nS_n^*F^*)$.

By adding the four terms in (7), we see that $\delta(FA_nA_n^*F^*)$ can be obtained by using $11n$ multiplications, $5n$ additions and 5 FFTs. We note that by (6),

$$\delta(FA_n^*F^*) = \delta(FA_nF^*)^* = [\Lambda_c + \delta(FS_nF^*)]^*,$$

where Λ_c and $\delta(FS_nF^*)$ are already computed in part (i) and (ii) above. Thus $\delta(FA_n^*F)$ can be computed in n additions. By (5), we see that $\widehat{C}_n e_0$ can be computed by an additional n multiplications and one FFT. Finally, by recalling that C_ne_0 and S_ne_0 are computed in $2n$ additions and $2n$ multiplications, we see that $\widehat{C}_n$ can be obtained in totally $8n$ additions, $14n$ multiplications and 6 FFTs. As remarked above, the last FFT can be saved because we only need to know the eigenvalues of $\widehat{C}_n$ but not its explicit form. Comparing with the algorithm proposed in Tyrtyshnikov [7] which requires 9 FFTs and $O(n)$ operations, we see that our method is more efficient.

References

[1] R. Chan and G. Strang, *Toeplitz Equations by Conjugate Gradients with Circulant Preconditioner*, SIAM J. Sci. Statist. Comput., 10 (1989), pp. 104–119.

[2] R. Chan, *Circulant Preconditioners for Hermitian Teoplitz Systems*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 542–550.

[3] R. Chan, X. Jin and M. Yeung, *The Spectra of Super-optimal Circulant Preconditioned Toeplitz Systems*. To appear .

[4] T. Chan, *An Optimal Circulant Preconditioner for Toeplitz Systems*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 766–771.

[5] P. Davis, *Circulant Matrices*, John Wiley & Sons, Inc., New York, 1979.

[6] G. Strang, *A Proposal for Toeplitz Matrix Calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.

[7] E. Tyrtyshnikov, *Optimal and Super-optimal Circulant Preconditioners*. To appear .