Zhou *Journal of Inequalities and Applications* (2022) 2022:126
https://doi.org/10.1186/s13660-022-02863-4

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

Properties and applications of a conjugate transform on Schatten classes

Weiqi Zhou[1*]

*Correspondence: zwq@xzit.edu.cn
[1]School of Mathematics and Statistics, Xuzhou University of Technology, Lishui Road 2, Yunlong District, Xuzhou, Jiangsu Province 221111, China

Abstract

We study a "conjugate" transform on matrix spaces. For Laurent/Toeplitz operators such a transform is a way of realizing the Hilbert transform on the torus. We establish its operator norm on Schatten classes and discuss the possibility of its boundedness upon permutations. Applications in the Rademacher–Menshov inequality and iterative methods are also included.

MSC: 15A04; 15A39; 15B05; 15B57; 42A50; 44A15

Keywords: Rademacher–Menshov; Triangular truncation; Hilbert transform; Conjugate transform; Gauss–Seidel

1 Motivation

We consider the following transform $\tilde{T}$ on matrix spaces:

$$\tilde{T}(H) = T \circ H,$$

where $\circ$ is the Hadamard product, $H \in \mathbb{C}^{n\times n}$, and

$$T = \begin{pmatrix} \ddots & & i \\ & 0 & \\ -i & & \ddots \end{pmatrix}. \tag{1}$$

If $H = L + L^*$ ($*$ denotes the adjoint) with L being strictly lower triangular, then

$$L = \frac{L+L^*}{2} + i \cdot \frac{L-L^*}{2i} = \frac{1}{2}H + \frac{i}{2}\tilde{T}(H),$$

thus $\tilde{T}$ simply takes the "real" part of L to its "imaginary" part, because of this it should be reasonable to call $\tilde{T}$ the conjugate transform.

Another good reason for such a name is the connection of $\tilde{T}$ to the Hilbert transform on the torus, which is defined as

$$f \mapsto \tilde{f}(\theta) = \frac{1}{2\pi}\text{p.v.}\int_0^{2\pi} f(t)\cot\left(\frac{\theta - t}{2}\right)dt,$$

(p.v. stands for Cauchy principal value). The Fourier series of f and $\tilde{f}$ differ by a sign depending on the frequency term, i.e., if $f(\theta) = \sum_{k\in\mathbb{Z}}\hat{f}(k)e^{ik\theta}$, then $\tilde{f}(\theta) = -i\sum_{k\in\mathbb{Z}}\operatorname{sgn}(k)\hat{f}(k)e^{ik\theta}$ [1, Chap. 6].

Take $f \in L^\infty(\mathbb{T})$, it induces a bounded multiplication operator on $L^2(\mathbb{T})$ by $g \mapsto fg$. We can expand f into Fourier series (recall that $L^\infty(\mathbb{T}) \subset L^2(\mathbb{T})$, and by the Carleson theorem the series also converges a.e. [2, 3]) and write it as a vector $(\dots,\hat{f}(-1),\hat{f}(0),\hat{f}(1),\dots)^T$. In this way the multiplication operator induced by f can be represented by a bi-infinite Toeplitz matrix F (i.e., a Laurent operator) with $F_{ij} = \hat{f}(k)$ if $j - i = k$ (alternatively see [4, Chap. 1] or [5, Chap. 3]). It then follows that the multiplication operator $g \mapsto \tilde{f}g$ can be represented by the matrix $\tilde{T}(F)$, thus $\tilde{T}$ on matrix forms of Laurent/Toeplitz operators is a way of realizing the Hilbert transform on $L^\infty(\mathbb{T})$ (see also [6] for a different perspective where $\tilde{T}$ is viewed as the bilinear Hilbert transform on Hankel operators).

Moreover, we have

$$\frac{1}{2}\big(f(\theta) + i\tilde{f}(\theta) - \hat{f}(0)\big) = \sum_{k\in\mathbb{N}}\hat{f}(k)e^{ik\theta}.$$

The right-hand side is called the Riesz projection of f, on Toeplitz matrices it corresponds to

$$\frac{1}{2}\big(A + i\tilde{T}(A) - \tilde{D}(A)\big) = \tilde{L}(A), \tag{2}$$

where $\tilde{D}$ is the diagonal projection that maps A to its main diagonal, and $\tilde{L}$ is the triangular truncation that maps A to its strict lower triangular part. Since $\tilde{D}$ is for many norms bounded, the boundedness of $\tilde{L}$ can then essentially be determined by inspecting $\tilde{T}$.

The truncation $\tilde{L}$ appears at various places in mathematics, for example, in numerical analysis, $\tilde{L}$ enters critically into the iteration matrix for the Gauss–Seidel method and the Kaczmarz method, the error reduction rate with respect to the spectral condition number can be estimated using the spectral operator norm of $\tilde{L}$ (see [7, 8]); In functional analysis, $\tilde{L}$ on finite dimensional spaces is the explicit form of the projection that maps a Schatten class to the subclass of Volterra operators in it (see [9, Chap. 3] or [10]); In harmonic analysis, the norm of majorant function in the Rademacher–Menshov inequality [11, 12] can be estimated by the norm $\tilde{L}$ (see [13]). Therefore, as simple as the form of $\tilde{T}$ (and $\tilde{L}$) is, its rich and profound background intrigues us to understand its behavior on $\mathbb{C}^{n\times n}$.

To our interest is the Schatten class S_p, which consists of compact operators whose singular values are in ℓ^p. S_p is a Banach space equipped with the ℓ^p norm of its singular values. S_1, S_2, S_∞ norms are nuclear, Hilbert–Schmidt, and spectral norms respectively. We use $\|\cdot\|_p$ to denote the S_p norm of a matrix, if $p = \infty$, then the subscript is omitted.

For integral operators, it is known that if their symbol belongs to particular mixed norm spaces $L^{p,q}$ (p, q are Hölder conjugates with $p \geq 2$), then they are in the Schatten class S_p (see [14–16]). On the other hand, the Hilbert transform is bounded on $L^p(\mathbb{T})$ for $1 < p < \infty$ (known as the Marcel–Riesz inequality [1, Chap. 6.17]) and unbounded on $L^1(\mathbb{T})$ (thus also unbounded on $L^\infty(\mathbb{T})$ by duality and its anti-symmetry [17, Theorem 102]), an explicit example of this unboundedness can be found in [18, p. 250].

Such insights suggest that $\tilde{T}$ acts on S_p the same way as the Hilbert transform behaves on L^p, which brings us to the main result of this paper:

Theorem 1

(i) *The operator norm* $\|\tilde{T}\|_\infty$ *of* $\tilde{T}$ *on* $\mathbb{C}^{n\times n}$ *with respect to the* S_∞ *norm is*

$$\|\tilde{T}\|_\infty = \frac{1}{n}\|T\|_1 = \frac{1}{n}\sum_{k=0}^{n-1}\left|\cot\frac{(2k+1)\pi}{2n}\right| \asymp \frac{2}{\pi}\ln n.$$

(ii) *The operator norm* $\|\tilde{T}\|_p$ *of* $\tilde{T}$ *on* $\mathbb{C}^{n\times n}$ *with respect to the* S_p *norm for* $2 \le p < \infty$ *satisfies* (*regardless of the dimension* n)

$$\|\tilde{T}\|_p \le 4p.$$

(iii) *The following holds regardless of the size of* A:

$$\sup_{\operatorname{rank}(A)=r}\frac{\|\tilde{T}(A)\|}{\|A\|} \le 4e\ln r.$$

(iv) *For any* $A \in \mathbb{C}^{n\times n}$, *there exist a permutation matrix* P *and a constant* C *independent of the dimension such that*

$$\left\|\tilde{T}(PAP^*)\right\| \le C\|A\|.$$

(v) *There is a constant* C *independent of the dimension and the choice of* $A \in \mathbb{C}^{n\times n}$ *such that*

$$\left\|\frac{1}{n!}\sum_P \tilde{T}(PAP^*)\right\| \le C\|A\|,$$

where the summation is taken over all possible permutation matrices P.

2 Preliminaries

If $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{C}^n$, then we write

$$D_x = \operatorname{diag}(x_1, x_2, \dots, x_n), \qquad P_x = xx^*,$$

in particular, one may verify that

$$P_x \circ A = D_x A D_x^*. \tag{3}$$

Let

$$\zeta = e^{\frac{\pi i}{n}}, \qquad \omega = e^{\frac{2\pi i}{n}} = \zeta^2, \tag{4}$$

and denote W as the Fourier matrix whose ijth entry is $W_{ij} = \omega^{(i-1)(j-1)}/\sqrt{n}$.

Lemma 1 *T can be diagonalized as*

$$T = D_\xi^* W^* D_\tau W D_\xi,$$

where

$$\xi = \left(1, \zeta, \zeta^2, \dots, \zeta^{n-1}\right)^T, \qquad \tau = (\tau_0, \tau_1, \dots, \tau_{n-1})^T,$$

with

$$\tau_k = \cot\frac{(2k+1)\pi}{2n}.$$

Proof It is easy to verify that $D_\xi T D_\xi^*$ is circulant, thus it can be diagonalized by W, the cotangent comes from further computation that

$$\tau_k = -i\sum_{j=1}^{n-1}(z_k)^j = -i\left(\frac{1-(z_k)^n}{1-z_k} - 1\right) = -i\left(\frac{1+z_k}{1-z_k}\right) = -i\left(\frac{z_k^{-\frac{1}{2}} + z_k^{\frac{1}{2}}}{z_k^{-\frac{1}{2}} - z_k^{\frac{1}{2}}}\right) = \cot\frac{(2k+1)\pi}{2n},$$

where $z_k = \zeta\omega^k$. □

Lemma 2 *Let H be a Hermitian matrix with vanishing main diagonal, if*

$$c_p = \sup_H \frac{\|\tilde{T}(H)\|_p}{\|H\|_p},$$

with the supreme taken over all such matrices, then

$$c_p \le p.$$

Proof The inequality is obvious for $p = 2$ since $c_2 = 1 < 2$, now suppose it holds for $p = k$, and we look at the case of $p = 2k$. Notice that the following holds:

$$H\tilde{T}(H) + \tilde{T}(H)H = -i(L+U)(L-U) - i(L-U)(L+U) = -2i\left(L^2 - U^2\right),$$

where L, U are respectively the strict lower and upper triangular part of H. It follows that

$$H^2 + \tilde{T}\left(H\tilde{T}(H) + \tilde{T}(H)H\right) = (L+U)^2 - 2\left(L^2 + U^2\right) = -(L-U)^2 = \left(\tilde{T}(H)\right)^2,$$

thus

$$\begin{aligned}
\left\|\tilde{T}(H)\right\|_{2k}^2 &= \left\|\left(\tilde{T}(H)\right)^2\right\|_k \\
&\le \left\|H^2\right\|_k + \left\|\tilde{T}\left(H\tilde{T}(H) + \tilde{T}(H)H\right)\right\|_k \\
&\le \|H\|_{2k}^2 + 2c_k\|H\|_{2k}\left\|\tilde{T}(H)\right\|_{2k},
\end{aligned}$$

i.e.,

$$c_{2k}^2 \le 1 + 2c_k c_{2k},$$

which we may solve and get

$$c_{2k} \le c_k + \sqrt{1 + c_k^2}.$$

By induction it then leads to

$$c_{2^n} \leq 2^n.$$

For other values of p, simply apply the Riesz–Thorin interpolation theorem. □

3 Proof of the main theorem

Proof

(i) By (3) and Lemma 1, we have

$$\|\tilde{T}(A)\| \leq \sum_{k=0}^{n-1} \|\tau_k D_{u_k} A D_{u_k}^*\| = \frac{1}{n}\sum_{k=0}^{n-1} |\tau_k| \|A\| = \frac{1}{n}\|T\|_1 \|A\|,$$

where u_k is the $k+1$st column in $D_z^* W^*$. The equality is attainable at, e.g.,

$$A = W^* D_{\operatorname{sgn}(\tau)} W,$$

where

$$\operatorname{sgn}(\tau) = \left(\operatorname{sgn}(\tau_0), \operatorname{sgn}(\tau_1), \ldots, \operatorname{sgn}(\tau_{n-1})\right)^T.$$

The asymptotic estimate follows by noticing that

$$\frac{\pi}{2n}\sum_{k=0}^{n-1}\left|\cot\frac{(2k+1)\pi}{2n}\right| \asymp \int_{\frac{\pi}{4n}}^{\frac{4n-1}{4n}} |\cot x|\, dx,$$

where the left-hand side can be viewed as a quadrature formula (e.g., middle point rule) for the integral in the right-hand side, which grows like $\ln n$.

(ii) Denote $\tilde{A} = A - D(A)$, then apply Lemma 2 to get

$$\begin{aligned}\|\tilde{T}(A)\|_p &= \|\tilde{T}(\tilde{A})\|_p \\ &\leq \frac{1}{2}\|\tilde{T}(\tilde{A}+\tilde{A}^*)\|_p + \frac{1}{2}\|\tilde{T}(\tilde{A}-\tilde{A}^*)\|_p \\ &\leq 2p\|\tilde{A}\|_p \leq 4p\|A\|_p.\end{aligned}$$

(iii) This is a direct consequence of (ii) since

$$\|\tilde{T}(A)\| \leq \|\tilde{T}(A)\|_p \leq 4p\|A\|_p \leq 4pr^{\frac{1}{p}}\|A\| \leq 4e\ln r\|A\|,$$

where the bound in the last inequality is attained at $p = \ln r$ (easily verifiable with elementary calculus).

(iv) The proof critically relies on the following celebrated paving conjecture (now a theorem) [19]:

Paving: For every ϵ with $1 > \epsilon > 0$, there exists a number γ_ϵ, which depends only on ϵ, such that for any $A \in \mathbb{C}^{n\times n}$ with vanishing main diagonal, one can partition the set $\{1, 2, \dots, n\}$ into γ_ϵ number of subsets $\Lambda_1, \Lambda_2, \dots, \Lambda_{\gamma_\epsilon}$ with the property that

$$\left\| Q_{\Lambda_i} A Q^*_{\Lambda_i} \right\| \le \epsilon \|A\|, \quad i = 1, 2, \dots, \gamma_\epsilon,$$

where Q_{Λ_i} is the orthogonal projection onto the space spanned by $\{\vec{e}_k\}_{k\in\Lambda_i}$ with $\vec{e}_k$ being the kth standard Euclidean basis vector.

The paving conjecture is an equivalent formulation of the Kadison–Singer problem [20], which was solved in [21]. It suffices to take γ_ϵ to be $(6/\epsilon)^4$ for real matrices and $(6/\epsilon)^8$ for complex matrices, see the exposition in [22].

Clearly, for our problem it suffices (since diagonal projections are bounded) to consider only matrices with vanishing main diagonals. The existence of such a permutation can then be established by induction, and we may take

$$C = \frac{2(\gamma_\epsilon - 1)}{1 - \epsilon}$$

for some properly chosen ϵ.

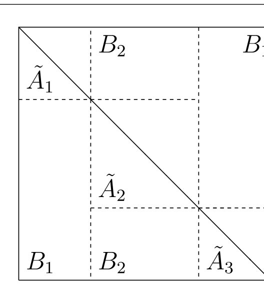
For $n = 2$, the statement is trivially true for, e.g., $\epsilon = 1/2$. Suppose it holds for all $n \le m$, and consider the case of $m + 1$. For a matrix A with vanishing main diagonal, we pave A to get the partition $\Lambda_1, \Lambda_2, \dots, \Lambda_{\gamma_\epsilon}$ and simultaneously permute (denote the permutation as σ) rows and columns of A so that $\{Q_{\Lambda_i} A Q^*_{\Lambda_i}\}_{i=1}^{\gamma_\epsilon}$ now appears as consecutive diagonal blocks of $P_\sigma A P^*_\sigma$. Denote $A_\sigma = P_\sigma A P^*_\sigma$.

Apply the induction assumption on each diagonal block $Q_{\Lambda_i} A_\sigma Q^*_{\Lambda_i}$ to obtain a permutation σ_i so that

$$\left\| \tilde{T}\left(P_{\sigma_i} Q_{\Lambda_i} A_\sigma Q^*_{\Lambda_i} P^*_{\sigma_i}\right) \right\| \le C \left\| Q_{\Lambda_i} A_\sigma Q^*_{\Lambda_i} \right\| \le C\epsilon \|A\|$$

holds. We combine these permutations $\sigma_1, \sigma_2, \dots, \sigma_{\gamma_\epsilon}$ and σ together to get a new matrix $\tilde{A}$. The strategy is best illustrated by Fig. 1. Each diagonal block of size $\Lambda_i \times \Lambda_i$ is denoted as $\tilde{A}_i$ in the above figure. Away from these diagonal blocks $\tilde{A}$ consists of $\gamma_\epsilon - 1$ number of matrices (denoted as B_i in the above figure), each of which consists of two rectangle matrices (both are submatrices of $\tilde{A}$) located in symmetric (with respect to the main diagonal) positions. Consequently, applying the induction assumption on the main diagonal blocks and the trivial estimate

Figure 1 The induction strategy



$\|B_i\| \le 2\|A\|$ elsewhere, we obtain

$$\big\|\tilde{T}(\tilde{A})\big\| \le \max_{1\le k\le\gamma_\epsilon} \big\|\tilde{T}(\tilde{A}_i)\big\| + \sum_{i=1}^{\gamma_\epsilon-1} \|B_i\| \le C\epsilon\|A\| + 2(\gamma_\epsilon - 1)\|\tilde{A}\| = C\|A\|.$$

(v) Consider the grand sum (i.e., the sum of all entries) of a matrix

$$\mathrm{gs}(A) = \sum_{j,k} A_{jk}. \tag{5}$$

It has a trivial upper bound

$$\big|\mathrm{gs}(A)\big| = \big|(A\vec{1},\vec{1})\big| \le n\|A\|, \tag{6}$$

where $\vec{1}$ is the all one vector. It is easy to see that for any matrix A we have

$$\sum_\sigma P_\sigma A P_\sigma^* = (n-2)!\big(\mathrm{gs}(A) - \mathrm{tr}(A)\big)E_0 + (n-1)!\,\mathrm{tr}(A)I,$$

where $E_0 = E - I$ with E being the all one matrix and I is the identity matrix, thus straightforward estimate shows

$$\frac{1}{n!}\bigg\|\sum_\sigma P_\sigma A P_\sigma^*\bigg\| \le c\|A\|,$$

with c being an absolute constant independent of n and A, since both $|\,\mathrm{gs}(A)|$ and $|\,\mathrm{tr}(A)|$ are trivially bounded by $n\|A\|$, while $\|E_0\| \le \|E\| + \|I\| \le n+1$. □

4 Applications

4.1 Optimal constants in Rademacher–Menshov inequality

The Rademacher–Menshov inequality [11, 12] states that if $\varphi = \{\varphi_k\}_{k\in\mathbb{N}}$ is an orthonormal system on some measure space (Ω,μ) and $a = \{a_k\}_{k\in\mathbb{N}} \in \ell^2$ is a scalar sequence, then

$$\|M_{a,\varphi,n}\|_{L^2} \le C\ln n\left(\sum_{k=1}^n |a_k|^2\right)^{\frac{1}{2}}, \tag{7}$$

where C is independent of a, φ, n and

$$M_{a,\varphi,n}(x) = \max_{m\le n}\left|\sum_{j=1}^m a_j\varphi_j(x)\right|$$

is often called the majorant function. With this inequality, one can further establish the Rademacher–Menshov theorem, i.e., if $\sum_{k=1}^\infty |a_k|^2\ln^2 k < \infty$, then $\sum_{k=1}^\infty a_k\varphi_k$ converges a.e. for all orthonormal systems $\{\varphi_k\}_{k\in\mathbb{N}}$. That boundedness of the majorant function implies a.e. convergence of the series is today a standard technique, see, e.g., the exposition in [23].

For convenience, let us denote

$$R_n = \frac{1}{\ln n}\sup_{a,\varphi}\frac{\|M_{a,\varphi,n}\|_{L^2}}{(\sum_{k=1}^n |a_k|^2)^{\frac{1}{2}}}. \tag{8}$$

For fixed n, R_n is the optimal constant in the right-hand side of (7) (while C in the right-hand side of the Rademacher–Menshov inequality (7) upper bounds R_n for all n). With the help of $\tilde{T}$, R_n can be estimated as follows.

Corollary 1 *For fixed n, the optimal constant R_n as defined in* (8) *in the Rademacher–Menshov inequality* (7) *satisfies $R_n \to \frac{1}{\pi}$ as $n \to \infty$.*

Proof Denote

$$L_n = \sup_{A\in\mathbb{C}^{n\times n}} \frac{\|\tilde{L}(A)\|}{\|A\|}, \qquad T_n = \sup_{A\in\mathbb{C}^{n\times n}} \frac{\|\tilde{T}(A)\|}{\|A\|}.$$

That $R_n \ln n = L_n$ can be justified in the following way (see also [13] for a different approach in probabilistic setting):

Let $\Lambda = \{\Lambda_j\}_{j=1}^n$ be a partition of Ω where each Λ_j is μ measurable. Compose the matrix $A^{(\Lambda)}$ whose elements are defined as

$$A_{ij}^{(\Lambda)} = \varphi_j|_{\Lambda_i},$$

then $A^{(\Lambda)}$ is a unitary linear map from $\mathbb{C}^n$ to $L^2(\Lambda_1) \oplus L^2(\Lambda_2) \oplus \cdots \oplus L^2(\Lambda_n)$ since if $\vec{a} = (a_1, a_2, \dots, a_n) \in \mathbb{C}^n$ and $f = a_1\varphi_1 + a_2\varphi_2 + \cdots + a_n\varphi_n$, then

$$\|f\|_{L^2}^2 = \left\|A^{(\Lambda)}\vec{a}\right\|_{L^2(\Lambda)}^2 = \|\vec{a}\|^2,$$

where $L^2(\Lambda)$ denotes $L^2(\Lambda_1) \oplus L^2(\Lambda_2) \oplus \cdots \oplus L^2(\Lambda_n)$. Now take

$$g_{\vec{a}} = \sum_{i=1}^{n} \sum_{j=1}^{j} a_j \varphi_j|_{\Lambda_i},$$

then we have

$$\|g_{\vec{a}}\|_{L^2}^2 = \left\|\tilde{L}\big(A^{(\Lambda)}\big)\vec{a}\right\|_{L^2(\Lambda)}^2,$$

consequently

$$L_n = \sup_{A^{(\Lambda)}} \frac{\|\tilde{L}(A^{(\Lambda)})\|}{\|A^{(\Lambda)}\|} = \sup_{\substack{A^{(\Lambda)}\\ \|\vec{a}\|=1}} \frac{\|\tilde{L}(A^{(\Lambda)})\vec{a}\|_{L^2(\Lambda)}^2}{\|A^{(\Lambda)}\|} = \sup_{\substack{\varphi\\ \|\vec{a}\|=1}} \|g_{\vec{a}}\|_{L^2} \le \sup_{\substack{\varphi\\ \|\vec{a}\|=1}} \|M_{a,\varphi,n}\|_{L^2} = R_n \ln n.$$

On the other hand, consider the following particular partition:

$$\tilde{\Lambda}_j = \left\{ x \in \Omega : \text{(i)} \left|\sum_{i=1}^{j} a_i\varphi_i(x)\right| > \left|\sum_{i=1}^{m} a_i\varphi_i(x)\right|, \forall m \le n; \right.$$
$$\left. \text{(ii) } j < j' \text{ if } \left|\sum_{i=1}^{j} a_i\varphi_i(x)\right| = \left|\sum_{i=1}^{j'} a_i\varphi_i(x)\right| \right\},$$

i.e., x belongs to $\tilde{\Lambda}_j$ if j is the smallest index where the sum $|\sum_{i=1}^{j} a_i\varphi_i(x)|$ attains the value of the majorant function $M_n(x)$ at x. Each $\tilde{\Lambda}_j$ is also measurable, since it is the pre-image of the measurable set range(M_n) under the function mapping $x \mapsto |\sum_{i=1}^{j} a_i\varphi_i(x)|$, thus we obtain that (with $\|a\| = 1$)

$$R_n \ln n = \|M_n(x)\|_{L^2}^2 = \sum_{j=1}^{n} \left\| \sum_{i=1}^{j} a_i\varphi_i(x) \right\|_{L^2(\tilde{\Lambda}_j)}^2 = \|\tilde{L}(A^{(\tilde{\Lambda})})\vec{a}\|_{L^2(\tilde{\Lambda})}^2 \le L_n,$$

together we get that $R_n = L_n$. It then easily follows from (2) and Theorem 1 (i) that

$$R_n = \frac{1}{\ln n} L_n \asymp \frac{1}{2\ln n} T_n \asymp \frac{1}{\pi}. \qquad \square$$

4.2 Ordering in Gauss–Seidel type methods

Let A be positive definite with diagonal D and strict lower triangular part L, then the error reduction matrix for applying the Gauss–Seidel method on a linear system $Ax = b$ is $Q = I - (D + L)^{-1}A$, thus with Theorem 1 (i) we can conclude that the error reduction rate per cycle is at least (see also [24])

$$1 - \frac{1}{c\kappa(A)\ln n}, \tag{9}$$

where $\kappa(A)$ is the spectral condition number of A and the constant c is independent of n, A and is approximately $1/\pi$.

A similar result holds for the Kaczmarz method [25], an alternating projection method also known as ART ([26]) whose randomized version has drawn much attention in recent years since [27]. Running the Kaczmarz method on $Ax = b$ is equivalent to running the Gauss-Seidel method implicitly on $AA^*y = b$ (see [28]). The Kaczmarz method converges even for rank deficient A and inconsistent systems (see [29]), thus with Theorem 1 (iii), the error reduction rate in (9) can be improved in the rank deficient case by replacing the $\ln n$ factor with the $\ln r$ factor, the same also holds for the Gauss–Seidel method on positive semi-definite matrices.

An often observed phenomena in reality is that rearranging the ordering of equations may (though need not) accelerate the error reduction. Theorem 1 (iv) and (v) provides an explanation: The linear system in natural ordering (given ordering) may converge slowly in bad cases where the $\ln n$ factor in (9) may be active, but by (iv) there exists some good ordering with which one can get rid of this $\ln n$ factor, while (v) shows that shuffling equations after each sweep will on average also remove it.

Acknowledgements
Not applicable.

Funding
Not applicable.

Availability of data and materials
Not applicable.

Declarations

Competing interests
The author declares that they have no competing interests.

Author contribution

The author read and approved the final manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 6 October 2021 Accepted: 2 September 2022 Published online: 24 September 2022

References

1. King, F.: Hilbert Transform (Vol 1 and 2). Cambridge University Press, Cambridge (2009)
2. Carleson, L.: On convergence and growth of partial sums of Fourier series. Acta Math. **116**(1), 135–157 (1966)
3. Hunt, R.: On the convergence of Fourier series, orthogonal expansions and their continuous analogues. In: Proceedings of the Conference at Edwardsville Ill, pp. 235–255 (1967)
4. Böttcher, A., Grudsky, S.: Toeplitz Matrices, Asymptotic Linear Algebra, and Functional Analysis. Springer, Berlin (2000)
5. Gohberg, I., Goldberg, S., Kaashoek, M.: Basic Classes of Linear Operators. Springer, Berlin (2012)
6. Gasch, J., Gilbert, J.: Triangularization of Hankel operators and the bilinear Hilbert transform. Contemp. Math. **247**, 235–248 (1999)
7. Oswald, P., Zhou, W.: Convergence analysis for Kaczmarz-type methods in a Hilbert space framework. Linear Algebra Appl. **478**, 131–161 (2015)
8. Oswald, P., Zhou, W.: Random reordering in SOR-type methods. Numer. Math. **135**(4), 1207–1220 (2017)
9. Gohberg, I., Krein, M.: Theory and Applications of Volterra Operators in Hilbert Space. Am. Math. Soc., Providence (1970)
10. Davidson, K.: Nest Algebras. Longman, Harlow (1988)
11. Rademacher, H.: Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen. Math. Ann. **87**(1–2), 112–138 (1922)
12. Meńshov, D.: Sur les séries de fonctions orthogonales I. Fundam. Math. **4**, 82–105 (1923)
13. Chobanyan, S., Levental, S., Salehi, H.: On the constant in Menshov–Rademacher inequality. J. Inequal. Appl. **2006**, 68969 (2006)
14. Russo, B.: On the Hausdorff–Young theorem for integral operators. Pac. J. Math. **68**(1), 241–253 (1977)
15. Goffeng, M.: Analytic formulas for the topological degree of non-smooth mappings: the odd-dimensional case. Adv. Math. **231**(1), 357–377 (2012)
16. Rozenblum, G., Ruzhansky, M., Suragan, D.: Isoperimetric inequalities for Schatten norms of Riesz potentials. J. Funct. Anal. **271**(1), 224–239 (2016)
17. Titchmarsh, E.: Introduction to the Theory of Fourier Integrals. Oxford University Press, London (1948)
18. Wheeden, R., Zygmund, A.: Measure and Integral. Dekker, New York (1977)
19. Anderson, J.: Restrictions and representations of states on C^* algebras. Trans. Am. Math. Soc. **249**(2), 303–329 (1979)
20. Kadison, R., Singer, I.: Extensions of pure states. Am. J. Math. **81**(2), 383–400 (1959)
21. Marcus, A., Spielman, D., Srivastava, N.: Interlacing families II: mixed characteristic polynomials and the Kadison–Singer problem. Ann. Math. **182**(1), 327–350 (2015)
22. Casazza, P., Tremain, J.: Consequences of the Marcus/Spielman/Srivastava solution of the Kadison–Singer problem. In: Trends in Applied Harmonic Analysis, pp. 191–214. Springer, Berlin (2016)
23. Meaney, C.: Remarks on the Rademacher-Menshov Theorem. Research Symposium Asymptotic Geometric Analysis, Harmonic Analysis and Related Topics (2007)
24. Oswald, P.: On the convergence rate of SOR: a worst case estimate. Computing **52**(3), 245–255 (1994)
25. Kaczmarz, S.: Angenäherte Auflösung von Systemen linearer Gleichungen. Bull. Int. Acad. Pol. Sci. Lett. **35**, 355–357 (1937)
26. Gordon, R., Bender, R., Herman, G.: Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. J. Theor. Biol. **29**(3), 471–481 (1970)
27. Strohmer, T., Vershynin, R.: A randomized Kaczmarz algorithm with exponential convergence. J. Fourier Anal. Appl. **15**(2), 262–278 (2009)
28. Hackbusch, W.: Iterative Solution of Large Sparse Systems of Equations. Springer, Berlin (1994)
29. Galantai, A.: Projectors and Projection Methods. Springer, Berlin (2004)