Rome-Moscow school of Matrix Methods and Applied Linear Algebra
Rome, Sept 19 - Oct 3, 2010, Moscow, Oct 10 - Oct 24, 2010 Carmine Di Fiore

In these notes the concepts of circulants, τ and Toeplitz matrices, Hessenberg algebras and displacement decompositions, spaces of class $\mathbb{V}$ and best least squares fits on such spaces, are introduced and investigated. As a consequence of the results presented, the choice of matrices involved in displacement decompositions, the choice of preconditioners in solving linear systems and the choice of Hessian approximations in quasi-Newton minimization methods, become possible in wider classes of low complexity matrix algebras.

The Fourier matrix, circulants, and fast discrete transforms

Consider the following $n \times n$ matrix

$$P_1 = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ 1 & & & & 0 \end{bmatrix}.$$

Let $\omega \in \mathbb{C}$. Note that

$$P_1 \begin{bmatrix} 1 \\ 1 \\ \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 1 \\ \\ 1 \end{bmatrix}, \; P_1 \begin{bmatrix} 1 \\ \omega \\ \\ \omega^{n-1} \end{bmatrix} = \begin{bmatrix} \omega \\ \omega^2 \\ \\ \omega^{n-1} \\ 1 \end{bmatrix} = \omega \begin{bmatrix} 1 \\ \omega \\ \\ \omega^{n-1} \end{bmatrix},$$

where the latter identity holds if $\omega^n = 1$. More in general, if $\omega^n = 1$, we have the following vectorial identities

$$P_1 \begin{bmatrix} 1 \\ \omega^j \\ \\ \omega^{(n-1)j} \end{bmatrix} = \begin{bmatrix} \omega^j \\ \\ \omega^{(n-1)j} \\ 1 \end{bmatrix} = \omega^j \begin{bmatrix} 1 \\ \omega^j \\ \\ \omega^{(n-1)j} \end{bmatrix}, \; j = 0, 1, \ldots, n-1,$$

or, equivalently, the following matrix identity

$$P_1 W = W D_{1\omega^{n-1}},$$

$$D_{1\omega^{n-1}} = \begin{bmatrix} 1 & & \\ & \omega & \\ & & \\ & & \omega^{n-1} \end{bmatrix}, \; W = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^j & \omega^{n-1} \\ & & & \\ 1 & \omega^{n-1} & \omega^{(n-1)j} & \omega^{(n-1)(n-1)} \end{bmatrix}.$$

Proposition. If $\omega^n = 1$ and if $\omega^j \neq 1$ for $0 < j < n$, then $W^* W = nI$.

proof: since $|\omega| = 1$, $\overline{\omega} = \omega^{-1}$, we have

$$\begin{aligned} [W^*W]_{ij} &= [\overline{W}W]_{ij} = \textstyle\sum_{k=1}^n [\overline{W}]_{ik}[W]_{kj} = \sum_{k=1}^n \overline{\omega}^{(i-1)(k-1)} \omega^{(k-1)(j-1)} \\ &= \textstyle\sum_{k=1}^n \omega^{(k-1)(j-i)} = \sum_{k=1}^n (\omega^{j-i})^{k-1}. \end{aligned}$$

Thus $[W^*W]_{ij} = n$ if $i = j$, and $[W^*W]_{ij} = \frac{1-(\omega^{j-i})^n}{1-\omega^{j-i}} = 0$ if $i \neq j$ (note that the assumption $\omega^j \neq 1$ for $0 < j < n$ is essential in order to make $1 - \omega^{j-i} \neq 0$).

By the result of the above Proposition, we can say that the following (symmetric) Fourier matrix

$$F = \frac{1}{\sqrt{n}} W$$

is unitary, i.e. $F^*F = I$.

Exercise. Prove that $F^2 = JP_1$ where J is the permutation matrix $J\mathbf{e}_k = \mathbf{e}_{n+1-k}$, $k = 1, \ldots, n$ (J is usually called anti-identity).

The matrix identity satisfied by P_1 and W can be of course rewritten in terms of F, $P_1 F = F D_{1\omega^{n-1}}$, thus we obtain the equality

$$P_1 = F D_{1\omega^{n-1}} F^*$$

which states that the Fourier matrix diagonalizes the matrix P_1, or, more precisely, that *the columns of the Fourier matrix form a system of n unitarily orthonormal eigenvectors for the matrix P_1 with corresponding eigenvalues $1, \omega, \ldots, \omega^{n-1}$.*

But if F diagonalizes P_1, then it diagonalizes all polynomials in P_1:

$$\begin{aligned} P_1^{k-1} &= F D_{1\omega^{n-1}}^{k-1} F^*, \\ \textstyle\sum_{k=1}^n a_k P_1^{k-1} &= F \textstyle\sum_{k=1}^n a_k D_{1\omega^{n-1}}^{k-1} F^* \\ &= F \begin{bmatrix} \sum_{k=1}^n a_k & & \\ & \sum_{k=1}^n a_k \omega^{k-1} & \\ & & \sum_{k=1}^n a_k \omega^{(n-1)(k-1)} \end{bmatrix} F^* \\ &= F d(W\mathbf{a}) F^* = \sqrt{n} F d(F\mathbf{a}) F^* \end{aligned}$$

where by $d(\mathbf{z})$ we mean the diagonal matrix whose diagonal entries are $z_1, z_2, \ldots, z_n$.

Let us investigate the matrices P_1^{k-1}, $k = 1, \ldots, n$, and the matrix $\sum_{k=1}^n a_k P_1^{k-1}$ in the case $n = 4$:

$$P_1^0 = I,\ P_1^1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix},\ P_1^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},\ P_1^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$P_1^4 = P_1^3 P_1 = P_1^T P_1 = I = P_1^0,$$

$$\sum_{k=1}^4 a_k P_1^{k-1} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_4 & a_1 & a_2 & a_3 \\ a_3 & a_4 & a_1 & a_2 \\ a_2 & a_3 & a_4 & a_1 \end{bmatrix} = \sqrt{4} F d(F\mathbf{a}) F^*,\ F = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & \omega^3 \\ 1 & \omega^2 & \omega^4 & \omega^6 \\ 1 & \omega^3 & \omega^6 & \omega^9 \end{bmatrix},$$

$$\omega^4 = 1,\ \omega^j \neq 1,\ 0 < j < 4\ (\omega = e^{\pm \mathbf{i} 2\pi/4}).$$

Note that, for n generic, we have the identities $\mathbf{e}_1^T P_1^{k-1} = \mathbf{e}_k^T$, $k = 1, \ldots, n$, and $P_1^n = I$ (prove them!). So, the set $C = \{p(P_1)\}$ of all polynomials in P_1

is spanned by the matrices $J_k = P_1^{k-1}$; the particular polynomial $\sum_{k=1}^n a_k J_k$ is simply denoted by $C(\mathbf{a})$. Note that $C(\mathbf{a})$ is the matrix of C with first row $\mathbf{a}^T$:

$$C(\mathbf{a}) = \sum_{k=1}^n a_k J_k = \begin{bmatrix} a_1 & a_2 & & a_{n-1} & a_n \\ a_n & a_1 & & & a_{n-1} \\ & & & & \\ a_3 & & & & a_2 \\ a_2 & a_3 & & a_n & a_1 \end{bmatrix} = F d(F^T\mathbf{a}) d(F^T\mathbf{e}_1)^{-1} F^{-1}.$$

C is known as the space of circulant matrices.

Exercise. (i) Repeat all, starting from the $n \times n$ matrix

$$P_{-1} = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ -1 & & & & 0 \end{bmatrix}$$

and arriving to the (-1)-circulant matrix whose first row is $\mathbf{a}^T$, $\mathbf{a} \in \mathbb{C}^n$:

$$C_{-1}(\mathbf{a}) = \begin{bmatrix} a_1 & a_2 & & a_{n-1} & a_n \\ -a_n & a_1 & & & a_{n-1} \\ & & & & \\ -a_3 & & & & a_2 \\ -a_2 & -a_3 & & -a_n & a_1 \end{bmatrix}.$$

(ii) Let T be a Toeplitz $n \times n$ matrix, i.e. $T = (t_{i-j})_{i,j=1}^n$, for some $t_k \in \mathbb{C}$. Show that T can be written as the sum of a circulant and of a (-1)-circulant, that is, $T = C(\mathbf{a}) + C_{-1}(\mathbf{b})$, $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$.

Why circulant matrices can be interesting in the applications of linear algebra? The main reason is in the fact that the matrix-vector product $C(\mathbf{a})\mathbf{z}$ can be computed in at most $O(n \log_2 n)$ arithmetic operations (whereas, usually, a matrix-vector product requires n^2 multiplications).

Proposition FFT. Given $\mathbf{z} \in \mathbb{C}^n$, the complexity of the matrix-vector product $F\mathbf{z}$ is at most $O(n \log_2 n)$. Such operation is called discrete Fourier transform (DFT) of $\mathbf{z}$. As a consequence, the matrix-vector product $C(\mathbf{a})\mathbf{z}$ is computable by two DFTs (after the preprocessing DFT $F\mathbf{a}$).

proof: since $\omega^{(i-1)(k-1)}$ is the (i,k) entry of W and z_k is the k entry of $\mathbf{z} \in \mathbb{C}^n$, we have

$$\begin{aligned} (W\mathbf{z})_i &= \textstyle\sum_{k=1}^n \omega^{(i-1)(k-1)} z_k = \sum_{j=1}^{n/2} \omega^{(i-1)(2j-2)} z_{2j-1} + \sum_{j=1}^{n/2} \omega^{(i-1)(2j-1)} z_{2j} \\ &= \textstyle\sum_{j=1}^{n/2} (\omega^2)^{(i-1)(j-1)} z_{2j-1} + \sum_{j=1}^{n/2} \omega^{(i-1)(2(j-1)+1)} z_{2j} \\ &= \textstyle\sum_{j=1}^{n/2} (\omega^2)^{(i-1)(j-1)} z_{2j-1} + \omega^{i-1} \sum_{j=1}^{n/2} (\omega^2)^{(i-1)(j-1)} z_{2j}. \end{aligned}$$

Note that ω is in fact a function of n, i.e. the right notation for ω should be ω_n. Then $\omega^2 = \omega_n^2$ is such that $(\omega_n^2)^{n/2} = 1$ and $(\omega_n^2)^i \neq 1$ $0 < i < n/2$; in other words $\omega_n^2 = \omega_{n/2}$. So, we have the identities

$$(W_n\mathbf{z})_i = \sum_{j=1}^{n/2} \omega_{n/2}^{(i-1)(j-1)} z_{2j-1} + \omega_n^{i-1} \sum_{j=1}^{n/2} \omega_{n/2}^{(i-1)(j-1)} z_{2j}, \; i = 1, 2, \ldots, n. \quad (?)$$

It follows that, for $i = 1, \ldots, \frac{n}{2}$,

$$(W_n\mathbf{z})_i = (W_{n/2} \begin{bmatrix} z_1 \\ z_3 \\ \\ z_{n-1} \end{bmatrix})_i + \omega_n^{i-1}(W_{n/2} \begin{bmatrix} z_2 \\ z_4 \\ \\ z_n \end{bmatrix})_i.$$

Moreover, by setting $i = \frac{n}{2} + k$, $k = 1, \ldots, \frac{n}{2}$, in (?), we obtain

$$\begin{aligned} (W_n\mathbf{z})_{\frac{n}{2}+k} &= \sum_{j=1}^{n/2} \omega_{n/2}^{\frac{n}{2}(j-1)} \omega_{n/2}^{(k-1)(j-1)} z_{2j-1} + \omega_n^{\frac{n}{2}} \omega_n^{k-1} \sum_{j=1}^{n/2} \omega_{n/2}^{\frac{n}{2}(j-1)} \omega_{n/2}^{(k-1)(j-1)} z_{2j} \\ &= \sum_{j=1}^{n/2} \omega_{n/2}^{(k-1)(j-1)} z_{2j-1} - \omega_n^{k-1} \sum_{j=1}^{n/2} \omega_{n/2}^{(k-1)(j-1)} z_{2j} \\ &= (W_{n/2} \begin{bmatrix} z_1 \\ z_3 \\ \\ z_{n-1} \end{bmatrix})_k - \omega_n^{k-1}(W_{n/2} \begin{bmatrix} z_2 \\ z_4 \\ \\ z_n \end{bmatrix})_k, \ k = 1, \ldots, \tfrac{n}{2}. \end{aligned}$$

($\omega_n^{\frac{n}{2}} = -1$; think $\omega = e^{\pm \mathbf{i}2\pi/n}$). Thus

$$W_n\mathbf{z} = \begin{bmatrix} I & D_{1\omega_n^{\frac{n}{2}-1}} \\ I & -D_{1\omega_n^{\frac{n}{2}-1}} \end{bmatrix} \begin{bmatrix} W_{n/2} & 0 \\ 0 & W_{n/2} \end{bmatrix} Q\mathbf{z},$$

$$D_{1\omega_n^{\frac{n}{2}-1}} = \begin{bmatrix} 1 & & \\ & \omega_n & \\ & & \omega_n^{\frac{n}{2}-1} \end{bmatrix}, Q = \begin{bmatrix} 1 & & & & & \\ 0 & 0 & 1 & & & \\ & & & & & \\ & & & & 1 & 0 \\ 0 & 1 & & & & \\ 0 & 0 & 0 & 1 & & \\ & & & & & \\ & & & & 0 & 1 \end{bmatrix}. \qquad (??)$$

If c_n denotes the complexity of the matrix-vector product $F_n\mathbf{z}$, then, by the previous formula,

$$c_n \leq 2c_{n/2} + rn, \ r \text{ constant.}$$

But this implies $c_n = O(n \log_2 n)$. The proof of the last assertion is left to the reader.

Of course, any time a $n \times n$ matrix U, well defined for all n, satisfies for n even an identity of the type

$$U_n = \begin{bmatrix} \text{sparse matrix} \end{bmatrix} \begin{bmatrix} U_{n/2} & 0 \\ 0 & U_{n/2} \end{bmatrix} \begin{bmatrix} \text{permutation matrix} \end{bmatrix},$$

the matrix-vector product $U_n\mathbf{z}$ can be computed in at most $O(n \log_2 n)$ arithmetic operations. The above identity is verified for at least 10 matrices U, the Fourier transform and its (-1) version, and the eight Hartley-type transforms. Note, however, that there are also other 16 discrete transforms of complexity $O(n \log_2 n)$, sine-type and the cosine-type transforms. See [],[].

Exercise G. Prove that the $n \times n$ matrix G defined by

$$G_{ij} = \frac{1}{\sqrt{n}}(\cos\frac{(2i+1)(2j+1)}{2n} + \sin\frac{(2i+1)(2j+1)}{2n}), \ i, j = 0, \ldots, n-1,$$

is symmetric, persymmetric, real, unitary, and satisfies the identity:

$$G_n = \frac{1}{\sqrt{2}} \begin{bmatrix} R_+ & R_- \\ -R_-J & R_+J \end{bmatrix} \begin{bmatrix} G_{n/2} & 0 \\ 0 & G_{n/2} \end{bmatrix} Q, \; R_\pm = D_c \pm D_s J,$$

(J $\frac{n}{2}\times\frac{n}{2}$ anti-identity) for some suitable $\frac{n}{2}\times\frac{n}{2}$ diagonal matrices D_c, D_s. Prove, moreover, that each row of G_n has at least a zero entry when $n = 2+4s$ (this is like to say, we will see, that the space $\{Gd(\mathbf{z})G : \mathbf{z} \in \mathbb{C}^n\}$ is not a h-space for such values of n); and that, instead, for all other n, $[G_n]_{1k} \neq 0 \; \forall\, k$ (i.e., for all other n, the space $\{Gd(\mathbf{z})G : \mathbf{z} \in \mathbb{C}^n\}$ is a 1-space).

Exercise. Prove that the space $C^S_{-1} + JC^{SK}_{-1}$, C^S_{-1} symmetric $n \times n$ (-1)-circulants, C^{SK}_{-1} skewsymmetric (-1)-circulants, is a commutative matrix algebra (a matrix A is skewsymmetric if $A^T = -A$).

The sine matrix and the (commutative) algebra of τ matrices

Consider the $n \times n$ matrix

$$P_0 + P_0^T = \begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ 0 & 1 & 0 & & \\ & & & & 1 \\ & & & 1 & 0 \end{bmatrix},$$

and set $J_1 = I$, and $J_2 = P_0 + P_0^T$. Note that $\mathbf{e}_1^T J_1 = \mathbf{e}_1^T$, $\mathbf{e}_1^T J_2 = \mathbf{e}_2^T$. Moreover, since

$$(P_0+P_0^T)^2 = \begin{bmatrix} 1 & 0 & 1 & & & & \\ 0 & 2 & 0 & 1 & & & \\ 1 & 0 & 2 & & & & \\ & 1 & & & & & \\ & & & & & & 1 \\ & & & & & 2 & 0 \\ & & & & & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & & & & \\ 0 & 1 & 0 & 1 & & & \\ 1 & 0 & 1 & & & & \\ & 1 & & & & & \\ & & & & & & 1 \\ & & & & & 1 & 0 \\ & & & & 1 & 0 & 0 \end{bmatrix} + I,$$

we have $\mathbf{e}_1^T((P_0 + P_0^T)^2 - I) = [0\,0\,1\,0 \cdots 0] = \mathbf{e}_3^T$. Set $J_3 = (P_0 + P_0^T)^2 - I = J_2(P_0 + P_0^T) - J_1$; then $\mathbf{e}_1^T J_3 = \mathbf{e}_3^T$.

More in general, set $J_{i+1} = J_i(P_0+P_0^T)-J_{i-1}$, $i = 2, 3, \ldots, n-1$. The matrix J_{i+1} is a polynomial in $P_0 + P_0^T$ of degree i with the property $\mathbf{e}_1^T J_{i+1} = \mathbf{e}_{i+1}^T$.

proof: assume $\mathbf{e}_1^T J_j = \mathbf{e}_j^T$, $j = 1, \ldots, i$; then

$$\mathbf{e}_1^T J_{i+1} = \mathbf{e}_1^T(J_i(P_0+P_0^T)-J_{i-1}) = (\mathbf{e}_i^T(P_0+P_0^T))-\mathbf{e}_{i-1}^T = (\mathbf{e}_{i-1}^T+\mathbf{e}_{i+1}^T)-\mathbf{e}_{i-1}^T = \mathbf{e}_{i+1}^T.$$

Since $J_1, J_2, \ldots, J_n$ are linearly independent, we can say that they span the set $\{p(P_0 + P_0^T)\}$ of all polynomials in the matrix $P_0 + P_0^T$ (use the Cailey-Hamilton theorem). We call such set τ. Note that the matrices of τ are determined once their first row is known; with the symbol $\tau(\mathbf{a})$ we denote the matrix of τ whose first row is $\mathbf{a}^T$, i.e. the matrix $\sum_k a_k J_k$.

Let us find a useful representation of τ and, in particular, of $\tau(\mathbf{a})$. First observe that the following vectorial equalities hold:

$$\begin{bmatrix} & 1 & & & \\ 1 & 0 & 1 & & \\ & 1 & & & \\ & & & & 1 \\ & & & 1 & \end{bmatrix} \begin{bmatrix} \sin\frac{j\pi}{n+1} \\ \sin\frac{2j\pi}{n+1} \\ \\ \sin\frac{nj\pi}{n+1} \end{bmatrix} = 2\cos\frac{j\pi}{n+1} \begin{bmatrix} \sin\frac{j\pi}{n+1} \\ \sin\frac{2j\pi}{n+1} \\ \\ \sin\frac{nj\pi}{n+1} \end{bmatrix}, \quad j = 1, \ldots, n.$$

Such n equalities can be rewritten as a simple matrix identity $(P_0 + P_0^T)S = SD$ where S is the matrix

$$S_{ij} = \sqrt{\frac{2}{n+1}} \sin\frac{ij\pi}{n+1}, \quad i, j = 1, \ldots, n,$$

and D is the diagonal matrix with diagonal entries $D_{jj} = 2\cos\frac{j\pi}{n+1}$. Note that the matrix S, called sine matrix, is real, symmetric and unitary (prove it!).

Remark. Let $F_{2(n+1)}$ be the Fourier matrix of order $2(n+1)$. Then the sine matrix S satisfies the following relation:

$$\big(I - F_{2(n+1)}^2\big)F_{2(n+1)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & S & 0 & -SJ \\ 0 & 0 & 0 & 0 \\ 0 & -JS & 0 & JSJ \end{bmatrix}$$

(note that $F_{2(n+1)}$ is a permutation matrix). As a consequence, a sine transform can be computed by performing a discrete Fourier transform.

So, the columns of the sine matrix S form a system of unitarily orthonormal eigenvectors for the matrix $P_0 + P_0^T$. In other words, the unitary matrix S diagonalizes $P_0 + P_0^T$ and, of course, diagonalizes any polynomial in $P_0 + P_0^T$, i.e. any τ matrix:

$$P_0 + P_0^T = SDS, \quad (P_0 + P_0^T)^k = SD^kS,$$
$$\tau = \{p(P_0 + P_0^T)\} = \{\textstyle\sum_{k=1}^n a_k(P_0 + P_0^T)^{k-1} : a_k \in \mathbb{C}\} = \{Sd(\mathbf{z})S : \mathbf{z} \in \mathbb{C}^n\}.$$

In particular, it is clear that the matrix of τ with first row $\mathbf{a}^T$ is

$$\tau(\mathbf{a}) = \sum_{k=1}^n a_k J_k = Sd(S^T\mathbf{a})d(S^T\mathbf{e}_1)^{-1}S^{-1}.$$

The latter formula states that matrix-vector products involving τ matrices have complexity at most $O(n\log_2 n)$.

Proposition PC. Given $X \in \mathbb{C}^{n\times n}$ generic, we have

$$\{p(X)\} \subset \{A \in \mathbb{C}^{n\times n} : AX = XA\},$$
$$\dim(\{p(X)\}) \le n \le \dim\{A \in \mathbb{C}^{n\times n} : AX = XA\}$$

and if one equality holds then the other equality holds too. So, X is non derogatory if and only if $\dim(\{p(X)\}) = n = \dim\{A \in \mathbb{C}^{n\times n} : AX = XA\}$.

The above Proposition suggests the following further representation of the space τ:

$$\tau = \{A \in \mathbb{C}^{n\times n} : A(P_0 + P_0^T) = (P_0 + P_0^T)A\}.$$

The fact that any matrix of τ must commute with $P_0 + P_0^T$ is equivalent to require that the following n^2 *cross-sum* conditions hold:

$$a_{i,j-1} + a_{i,j+1} = a_{i-1,j} + a_{i+1,j},\ \ i,j = 1,\dots,n$$

where we have set $a_{0,j} = a_{n+1,j} = a_{i,0} = a_{i,n+1} = 0$, $i,j = 1,\dots,n$. We can use such conditions in order to write down the generic τ matrix whose first row is $[a_1\, a_2\, \cdots\, a_n]$. For example, for $n = 4$, we can say that

$$\tau(\mathbf{a}) = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_2 & a_1 + a_3 & a_2 + a_4 & a_3 \\ a_3 & a_2 + a_4 & a_1 + a_3 & a_2 \\ a_4 & a_3 & a_2 & a_1 \end{bmatrix},\ J_1 = I,\ J_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$J_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},\ J_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},\ J_2 - J_4 = \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{bmatrix},$$

and so on.

Exercise. Prove that for n even the matrix J_2 is invertible, and, if possible, compute the inverse.

solution: We know that if J_2 is invertible, then $J_2^{-1} \in \tau$ (from the fact that J_2 commutes with $P_0 + P_0^T$ it follows that also J_2^{-1} commutes with $P_0 + P_0^T$!). Thus $J_2^{-1} = \tau(\mathbf{z})$ for some $\mathbf{z} \in \mathbb{C}^n$. Note that the matrix identity $\tau(\mathbf{z})J_2 = I$ is equivalent to the vectorial identity $\mathbf{z}^T J_2 = \mathbf{e}_1^T$. So, for example, for $n = 4$ we have the condition

$$[z_1\ z_2\ z_3\ z_4] \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = [1\ 0\ 0\ 0],$$

which yields $z_1 = 0$, $z_2 = 1$, $z_3 = 0$, $z_4 = -1$, and thus

$$J_2^{-1} = \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{bmatrix} = J_3 - J_4.$$

The proof for $n = 6, 8, \dots$ is left to the reader.

Exercise Ttau. Let T be a symmetric Toeplitz $n \times n$ matrix, i.e. $T = (t_{|i-j|})_{i,j=1}^n$, for some $t_k \in \mathbb{C}$. Show that $T = A + B$ where A is a τ matrix of order n and

$$B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & R & 0 \\ 0 & 0 & 0 \end{bmatrix},\ R \in \tau,\ R \in \mathbb{C}^{(n-2)\times(n-2)}.$$

Exercise. Write down rank one τ matrices (try first $n = 2$, $n = 3$, $n = 4$, $n = 5$, $n = 6$, $\dots$)

Hessenberg algebras

Let X be a lower Hessenberg 3×3 matrix

$$X = \begin{bmatrix} a_{11} & b_1 & 0 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

We now show that the space of all polynomials in X is spanned by three matrices J_1, J_2, J_3 such that $\mathbf{e}_1^T J_k = \mathbf{e}_k^T$, $k = 1, 2, 3$, provided that $X_{ii+1} \neq 0$, $i = 1, 2$.

As J_1 we take the identity, $J_1 = X^0 = I$. Let us define J_2:

$$X - a_{11}I = \begin{bmatrix} 0 & b_1 & 0 \\ a_{21} & a_{22} - a_{11} & b_2 \\ a_{31} & a_{32} & a_{33} - a_{11} \end{bmatrix},$$

$b_1 \neq 0 \Rightarrow$

$$J_2 = \frac{1}{b_1}(X - a_{11}I) = \begin{bmatrix} 0 & 1 & 0 \\ a_{21}/b_1 & (a_{22} - a_{11})/b_1 & b_2/b_1 \\ a_{31}/b_1 & a_{32}/b_1 & (a_{33} - a_{11})/b_1 \end{bmatrix}.$$

Note that $\mathbf{e}_1^T J_2 = \mathbf{e}_2^T$. Then, let us define J_3:

$$J_2^2 = \begin{bmatrix} a_{22}/b_1 & (a_{22} - a_{11})/b_1 & b_2/b_1 \\ & & \\ & & \end{bmatrix},$$

$b_2 \neq 0 \Rightarrow$

$$J_3 = \frac{b_1}{b_2}(J_2^2 - \frac{a_{22}}{b_1}I - \frac{a_{22} - a_{11}}{b_1}J_2).$$

Note that $\mathbf{e}_1^T J_3 = \mathbf{e}_3^T$. Finally, Cailey-Hamilton theorem yields the thesis, $\{p(X)\} = \text{Span}\,\{J_1, J_2, J_3\}$.

The following Proposition generalizes to a generic n the above remarks. For a detailed proof see [].

Proposition. Let X be a lower Hessenberg $n \times n$ matrix. Then the space H_X of all polynomials in X

$$H_X = \{\sum_{k=1}^{n} \alpha_k X^{k-1} : \alpha_k \in \mathbb{C}\}$$

is spanned by n matrices $J_1, \ldots, J_n$ such that $\mathbf{e}_1^T J_k = \mathbf{e}_k^T$, $k = 1, \ldots, n$, provided that $X_{ii+1} \neq 0$, $i = 1, \ldots, n-1$; in such case H_X is called Hessenberg algebra, and for any $\mathbf{a} \in \mathbb{C}^n$ there is a unique matrix in H_X with first row $\mathbf{a}^T$ which is denoted by $H_X(\mathbf{a})$, i.e. $H_X(\mathbf{a}) = \sum_k a_k J_k$.

Of course, by Proposition PC, any Hessenberg algebra H_X admits also the following representation

$$H_X = \{A \in \mathbb{C}^{n \times n} : AX = XA\}.$$

Until now we have seen two examples of Hessenberg algebras, ξ-circulants $\xi \neq 0$ ($X = P_\xi$) and tau matrices ($X = P_0 + P_0^T$). Both can be simultaneously diagonalized by a suitable matrix. An example of Hessenberg algebra whose matrices cannot be simultaneously diagonalized is H_{P_0}, the space of all upper triangular Toeplitz matrices. The matrix $H_{P_0}(\mathbf{a})$ is displayed here below

$$H_{P_0}(\mathbf{a}) = \begin{bmatrix} a_1 & a_2 & & & a_n \\ & a_1 & a_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & a_2 \\ & & & & a_1 \end{bmatrix}.$$

Even the matrix P_0 (i.e. the matrix generating the space) is not diagonalizable.

Note that when $X = P_0$ or, more in general, when $X = P_\xi$, the matrices J_k are simply the powers of X, i.e. $J_k = X^{k-1}$. For example, for $n = 3$

$$J_1 = I,\ J_2 = X = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \xi & 0 & 0 \end{bmatrix},\ J_3 = X^2 = \begin{bmatrix} 0 & 0 & 1 \\ \xi & 0 & 0 \\ 0 & \xi & 0 \end{bmatrix}.$$

Hessenberg algebras make up a subclass of commutative matrix algebras of the class of 1-spaces defined here below

Definition. $\mathcal{L} \subset \mathbb{C}^{n\times n}$ is said to be a 1-space if $\mathcal{L} = \text{Span}\,\{J_1, \dots, n\}$ with J_k such that $\mathbf{e}_1^T J_k = \mathbf{e}_k^T$, $k = 1, \dots, n$.

An example of 1-space $\mathcal{L}$ which is not a Hessenberg algebra is the following

$$\mathcal{L} = \{ \begin{bmatrix} A & JB \\ JB & A \end{bmatrix} : A, B\ \frac{n}{2} \times \frac{n}{2} \text{ circulants}\}.$$

One can easily prove that $\mathcal{L}$ is a non commutative matrix algebra. An example of 1-space which is not a matrix algebra is the set of all $n\times n$ symmetric Toeplitz matrices (see the next section).

Toeplitz linear systems and displacement decompositions

A $n \times n$ Toeplitz matrix is a matrix of the form $T = (t_{i-j})_{i,j=1}^n$. In applications often one has to solve Toeplitz linear systems $T\mathbf{x} = \mathbf{b}$.

For example, here below is a 3×3 Toeplitz matrix:

$$T = \begin{bmatrix} t_0 & t_{-1} & t_{-2} \\ t_1 & t_0 & t_{-1} \\ t_2 & t_1 & t_0 \end{bmatrix}.$$

An example of Toeplitz matrix T is the coefficient matrix of the linear system arising when solving, by finite differences or by finite elements, the boundary value differential problem $-u^{"} = f$, $u(a) = \alpha$, $u(b) = \beta$. In such case T is symmetric and its first row is $[2\ -1\ 0 \cdots 0]$. Another example, important is applied probability, is $T = (t^{|i-j|})_{i,j=1}^n$ with $|t|$ ($t \in \mathbb{C}$) less than 1.

Exercise. The vector space of all $n \times n$ *symmetric* Toeplitz matrices is a 1-space, being $(t_{|i-j|})_{i,j=1}^n$ equal to the sum $\sum_{k=1}^n t_{k-1} J_k$ where the J_k are the symmetric Toeplitz matrices with first row $\mathbf{e}_k^T$. Prove that such 1-space is not a matrix algebra.

In the framework of displacement theory it is possible to obtain some decompositions of T^{-1} involving matrices from Hessenberg algebras (or from more general commutative h-spaces) of the type

$$T^{-1} = \sum_{k=1}^{\alpha} M_k N_k, \ \ 2 \le \alpha \le 4.$$

Usually, the matrices M_k and N_k, appearing in such formulas, can be multiplied by a vector in $O(n \log_2 n)$ arithmetic operations; thus fast direct solvers of Toeplitz linear system naturally arise. Here below there is one example of such formulas:

$$T^{-1} = L_1 U_1 + L_2 U_2. \qquad (GS)$$

The L_j and U_j are suitable lower and upper triangular Toeplitz matrices, i.e. elements or transposed of elements from the Hessenberg algebra H_{P_0}.

Remark. By the Gohberg-Semencul formula (GS), if the L_j and U_j are known (a way to obtain them is indicated in []), then the matrix-vector product $T^{-1}\mathbf{b}$ can be computed in at most $O(n \log_2 n)$ arithmetic operations. That is, assuming preprocessing on T, the complexity of the problem of solving any Toeplitz system $T\mathbf{x} = \mathbf{b}$ is at most $O(n \log_2 n)$.

proof: it is enough to prove that any Toeplitz matrix (in particular the triangular ones) can be multiplied by a vector by means of a finite number of discrete Fourier transforms. The latter result is immediate if we observe that any $n \times n$ Toeplitx matrix can be embedded into a $(2n+k) \times (2n+k)$, $k \ge 0$, circulant matrix; for example, if $n = 3$ we have

$$T = \begin{bmatrix} t_0 & t_{-1} & t_{-2} \\ t_1 & t_0 & t_{-1} \\ t_2 & t_1 & t_0 \end{bmatrix}, \ C = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & & t_2 & t_1 \\ t_1 & t_0 & t_{-1} & t_{-2} & & t_2 \\ t_2 & t_1 & t_0 & t_{-1} & t_{-2} & \\ & t_2 & t_1 & t_0 & t_{-1} & t_{-2} \\ t_{-2} & & t_2 & t_1 & t_0 & t_{-1} \\ t_{-1} & t_{-2} & & t_2 & t_1 & t_0 \end{bmatrix}$$

$(k = 0)$. It is clear that the vector $T \cdot \mathbf{z}$ is the first part of the vector $C \begin{bmatrix} \mathbf{z} \\ \mathbf{0} \end{bmatrix}$.

The above representation (GS) for the inverse of a Toeplitz matrix, which can be very useful in order to solve Toeplitz linear systems efficiently [], follows from the *displacement decomposition* formula stated in the following theorem

Theorem DD. Let X be a lower Hessenberg $n \times n$ matrix. Assume that $X_{ij} = X_{i+1,j+1}$, $i, j = 1, \dots, n-1$ (X has Toeplitz structure), and that $b = X_{ii+1} \neq 0$, and consider the (commutative) Hessenberg algebra H_X generated by X (note that H_X is a 1-space).

Assume that $A \in \mathbb{C}^{n \times n}$ is such that $AX - XA = \sum_{m=1}^{\alpha} \mathbf{x}_m \mathbf{y}_m^T$. Then

$$bA = -\sum_{m=1}^{\alpha} H_{P_0}(\tilde{\mathbf{x}}_m)^T H_X(\mathbf{y}_m) + bH_X(A^T \mathbf{e}_1) \qquad (DD)$$

where, for $\mathbf{z} \in \mathbb{C}^n$, $H_{P_0}(\mathbf{z})$ is the upper triangular Toeplitz matrix with first row $\mathbf{z}^T$, $H_X(\mathbf{z})$ is the matrix of H_X with first row $\mathbf{z}^T$, and $\tilde{\mathbf{z}}$ is the vector $[0 \, z_1 \, \cdots \, z_{n-1}]^T$.

Note: Besides DD several other displacement decompositions hold, which can be general like DD, i.e. representing generic matrices A, or specialized for centrosymmetric A (see []). Such decompositions yield formulas for the inverses of Toeplitz, Toeplitz plus Hankel, and Toeplitz plus Hankel-like matrices useful in order to solve Toeplitz plus Hankel-like linear systems. Recall that a Hankel matrix is nothing else a matrix of the form JT where T is Toeplitz (the well known Hilbert matrix is an example of Hankel matrix).

In order to prove Theorem DD, the following Lemma is fundamental.

Lemma []. Let $\mathcal{L}$ be a commutative 1-space of $n \times n$ matrices, i.e. $\mathcal{L} = \{\sum_k \alpha_k J_k : \alpha_k \in \mathbb{C}\}$, with $J_k \in \mathbb{C}^{n\times n}$ such that $\mathbf{e}_1^T J_k = \mathbf{e}_k^T$, $J_k J_s = J_s J_k$. Let X be an element of $\mathcal{L}$, and assume that $A \in \mathbb{C}^{n\times n}$ is such that $AX - XA = \mathbf{x}\mathbf{y}^T$. Then $\mathbf{x}^T\mathcal{L}(\mathbf{y})^T = \mathbf{0}^T$.

proof: note that the equality $J_k J_s = J_s J_k$ implies $\mathbf{e}_1^T J_k J_s = \mathbf{e}_1^T J_s J_k$, $\mathbf{e}_k^T J_s = \mathbf{e}_s^T J_k$ $\forall\, s, k$, thus

$$\begin{aligned}
\mathbf{x}^T\mathcal{L}(\mathbf{y})^T\mathbf{e}_r &= \mathbf{x}^T(\textstyle\sum_k y_k J_k)^T\mathbf{e}_r = \mathbf{x}^T\sum_k y_k J_k^T\mathbf{e}_r \\
&= \mathbf{x}^T\textstyle\sum_k y_k J_r^T\mathbf{e}_k = \mathbf{x}^T J_r^T\sum_k y_k\mathbf{e}_k \\
&= \mathbf{x}^T J_r^T\mathbf{y} = \textstyle\sum_{i,j} x_i y_j [J_r^T]_{ij} \\
&= \textstyle\sum_{i,j} x_i y_j [J_r]_{ji} = \sum_{i,j}[AX - XA]_{ij}[J_r]_{ji} \\
&= \textstyle\sum_i [(AX - XA)J_r]_{ii} = \sum_i [(AJ_r)X - X(AJ_r)]_{ii} \\
&= \operatorname{tr}((AJ_r)X) - \operatorname{tr}(X(AJ_r)) = 0
\end{aligned}$$

(recall that the two matrices MN and NM, $M, N \in \mathbb{C}^{n\times n}$, have the same characteristic polynomial, even if (in case $\det(M) = \det(N) = 0$) MN and NM might be not similar each other).

We now report a draft of the proof of Theorem DD (for a more detailed proof see []). In order to obtain the equality (DD), which is of the type

$$bA = E + bH_X(A^T\mathbf{e}_1),$$

it is enough to prove that

$$EX - XE = (bA)X - X(bA), \qquad (***)$$

and to observe that the first row of E is null. In fact, the above equality implies $(bA - E)X - X(bA - E) = 0$, and thus $bA - E \in H_X$. The Lemma, applied for $\mathcal{L} = H_X$, is fundamental in proving (***).

A matrix algebra which is not a 1*-space: the class of spaces in* $\mathbb{V}$

Remember that a matrix A is said symmetric if $A^T = A$ ($a_{ji} = a_{ij}$), skewsymmetric if $A^T = -A$ ($a_{ji} = -a_{ij}$) and persymmetric if $A^T = JAJ$ ($a_{ji} = a_{n+1-i,n+1-j}$).

Consider a 6×6 symmetric (-1)-circulant matrix $A \in C_{-1}^S$,

$$A = \begin{bmatrix} a_0 & a_1 & a_2 & 0 & -a_2 & -a_1 \\ a_1 & a_0 & a_1 & a_2 & 0 & -a_2 \\ a_2 & a_1 & a_0 & a_1 & a_2 & 0 \\ 0 & a_2 & a_1 & a_0 & a_1 & a_2 \\ -a_2 & 0 & a_2 & a_1 & a_0 & a_1 \\ -a_1 & -a_2 & 0 & a_2 & a_1 & a_0 \end{bmatrix},$$

a 6×6 skewsymmetric (-1)-circulant matrix $B \in C_{-1}^{SK}$, and the matrix JB,

$$B = \begin{bmatrix} 0 & b_1 & b_2 & b_3 & b_2 & b_1 \\ -b_1 & 0 & b_1 & b_2 & b_3 & b_2 \\ -b_2 & -b_1 & 0 & b_1 & b_2 & b_3 \\ -b_3 & -b_2 & -b_1 & 0 & b_1 & b_2 \\ -b_2 & -b_3 & -b_2 & -b_1 & 0 & b_1 \\ -b_1 & -b_2 & -b_3 & -b_2 & -b_1 & 0 \end{bmatrix}, \; JB = \begin{bmatrix} -b_1 & -b_2 & -b_3 & -b_2 & -b_1 & 0 \\ -b_2 & -b_3 & -b_2 & -b_1 & 0 & b_1 \\ -b_3 & -b_2 & -b_1 & 0 & b_1 & b_2 \\ -b_2 & -b_1 & 0 & b_1 & b_2 & b_3 \\ -b_1 & 0 & b_1 & b_2 & b_3 & b_2 \\ 0 & b_1 & b_2 & b_3 & b_2 & b_1 \end{bmatrix}.$$

The vector space γ of all matrices of the type $A + JB$ has dimension equal to 6 (recall that $\dim(A + JB) = \dim(A) + \dim(JB) - \dim(A \cap JB)$).

We now show that there is not a basis $\{J_k\}$ for γ such that $\mathbf{e}_1^T J_k = \mathbf{e}_k^T$, $k = 1, 2, 3, 4, 5, 6$, i.e. γ is not a 1-space. Note that

$$\mathbf{e}_1^T(A + JB) = [(a_0 - b_1)(a_1 - b_2)(a_2 - b_3)(-b_2)(-a_2 - b_1)(-a_1)],$$

so, the equality $\mathbf{e}_1^T(A + JB) = \mathbf{e}_2^T$ is satisfied if and only if

$$a_0 - b_1 = 0,\ a_1 - b_2 = 1,\ a_2 - b_3 = 0,\ -b_2 = 0,\ -a_2 - b_1 = 0,\ -a_1 = 0.$$

Since we have both the conditions $b_2 = 0$ and $b_2 = -1$, a matrix $J_2 \in \gamma$ such that $\mathbf{e}_1^T J_2 = \mathbf{e}_2^T$ cannot exist.

However, there exists a basis $\{J_k\}$ of γ satisfying the equalities $(\mathbf{e}_1 + \mathbf{e}_6)^T J_k = \mathbf{e}_k^T$, $k = 1, 2, 3, 4, 5, 6$. For example, a matrix $J_2 \in \gamma$ with the property $(\mathbf{e}_1 + \mathbf{e}_6)^T J_2 = \mathbf{e}_2^T$ is obtained as follows. Note that

$$\mathbf{e}_6^T(A + JB) = [(-a_1)(-a_2 + b_1)(b_2)(a_2 + b_3)(a_1 + b_2)(a_0 + b_1)],$$

so, for the sum of the first and of the sixth rows of $A + JB$, we obtain the formula

$$(\mathbf{e}_1 + \mathbf{e}_6)^T(A + JB)$$
$$= [(a_0 - b_1 - a_1)(a_1 - b_2 - a_2 + b_1)(a_2 - b_3 + b_2)(-b_2 + a_2 + b_3)(-a_2 - b_1 + a_1 + b_2)(-a_1 + a_0 + b_1)].$$

Thus the condition $(\mathbf{e}_1 + \mathbf{e}_6)^T(A + JB) = \mathbf{e}_2^T$ is satisfied if and only if the following system of equations has solution

$$\begin{array}{l} a_0 - b_1 - a_1 = 0, \quad a_1 - b_2 - a_2 + b_1 = 1, \\ a_2 - b_3 + b_2 = 0, \quad -b_2 + a_2 + b_3 = 0, \\ -a_2 - b_1 + a_1 + b_2 = 0, \quad -a_1 + a_0 + b_1 = 0, \end{array}$$

and such system has the unique solution $a_0 = a_1 = \frac{1}{2}$, $a_2 = 0$, $b_2 = b_3 = -\frac{1}{2}$, $b_1 = 0$. The matrix $J_2 \in \gamma$ such that $(\mathbf{e}_1 + \mathbf{e}_6)^T J_2 = \mathbf{e}_2^T$ is displayed here below:

$$J_2 = \begin{bmatrix} \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{2} \\ 1 & 1 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}.$$

Analogously, one obtains the other J_k such that $(\mathbf{e}_1+\mathbf{e}_6)^T J_k = \mathbf{e}_k^T$:

$$J_1 = \begin{bmatrix} 1 & \frac12 & \frac12 & \frac12 & \frac12 & 0 \\ \frac12 & 1 & \frac12 & \frac12 & 0 & -\frac12 \\ \frac12 & \frac12 & 1 & 0 & -\frac12 & -\frac12 \\ \frac12 & \frac12 & 0 & 0 & -\frac12 & -\frac12 \\ \frac12 & 0 & -\frac12 & -\frac12 & 0 & -\frac12 \\ 0 & -\frac12 & -\frac12 & -\frac12 & -\frac12 & 0 \end{bmatrix}, \; J_3 = \begin{bmatrix} \frac12 & \frac12 & 1 & 0 & -\frac12 & -\frac12 \\ \frac12 & 1 & \frac12 & \frac12 & 0 & -\frac12 \\ 1 & \frac12 & \frac12 & \frac12 & \frac12 & 0 \\ 0 & \frac12 & \frac12 & \frac12 & \frac12 & 0 \\ -\frac12 & 0 & \frac12 & \frac12 & 0 & \frac12 \\ -\frac12 & -\frac12 & 0 & 0 & \frac12 & \frac12 \end{bmatrix},$$

($a_1=a_2=0$, $a_0=\frac12$, $b_2=b_3=-\frac12$, $b_1=-\frac12$), ($a_0=a_1=a_2=\frac12$, $b_1=b_2=0$, $b_3=-\frac12$),

$$J_4 = \begin{bmatrix} \frac12 & \frac12 & 0 & 0 & -\frac12 & -\frac12 \\ \frac12 & 0 & \frac12 & \frac12 & 0 & -\frac12 \\ 0 & \frac12 & \frac12 & \frac12 & \frac12 & 0 \\ 0 & \frac12 & \frac12 & \frac12 & \frac12 & 1 \\ -\frac12 & 0 & \frac12 & \frac12 & 1 & \frac12 \\ -\frac12 & -\frac12 & 0 & 1 & \frac12 & \frac12 \end{bmatrix}, \; J_5 = \begin{bmatrix} \frac12 & 0 & -\frac12 & -\frac12 & 0 & -\frac12 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac12 & 0 & \frac12 & \frac12 & 0 & \frac12 \\ -\frac12 & 0 & \frac12 & \frac12 & 1 & \frac12 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ -\frac12 & 0 & \frac12 & \frac12 & 1 & \frac12 \end{bmatrix},$$

($a_0=a_1=a_2=\frac12$, $b_1=b_2=0$, $b_3=\frac12$), ($a_0=a_1=\frac12$, $a_2=0$, $b_2=b_3=\frac12$, $b_1=0$),

$$J_6 = \begin{bmatrix} 0 & -\frac12 & -\frac12 & -\frac12 & -\frac12 & 0 \\ -\frac12 & 0 & -\frac12 & -\frac12 & 0 & \frac12 \\ -\frac12 & -\frac12 & 0 & 0 & \frac12 & \frac12 \\ -\frac12 & -\frac12 & 0 & 1 & \frac12 & \frac12 \\ -\frac12 & 0 & \frac12 & \frac12 & 1 & \frac12 \\ 0 & \frac12 & \frac12 & \frac12 & \frac12 & 1 \end{bmatrix}.$$

($a_1=a_2=0$, $a_0=\frac12$, $b_1=b_2=b_3=\frac12$).

Of course $\gamma=$ Span $\{J_1,J_2,J_3,J_4,J_5,J_6\}$ (the J_k are linearly independent!). The matrix $\sum_k a_k J_k$ is denoted by $\gamma(\mathbf{a})$. Note that $(\mathbf{e}_1+\mathbf{e}_6)^T\gamma(\mathbf{a})=\mathbf{a}^T$, so $\gamma(\mathbf{a})$ is called the matrix of γ whose $(\mathbf{e}_1+\mathbf{e}_6)$-row is $\mathbf{a}^T$.

More in general, one can easily prove that the set $\gamma = C_{-1}^S + JC_{-1}^{SK}$, C_{-1}^S $n\times n$ symmetric (-1)-circulants, C_{-1}^{SK} $n\times n$ skewsymmetric (-1)-circulants, is a vector space of dimension n, and is a commutative matrix algebra. Moreover, it is a 1-space if and only if n is one of the integers $\{3,4,5,7,8,9,11,12,13,15,\ldots\}$; for the remaining values of n, i.e. for $n=2+4s$ $s\in\mathbb{Z}$, no row of a matrix $A+JB$ of γ determines $A+JB$, that is, there is no index h for which there exists a basis $\{J_k\}$ of γ with the property $\mathbf{e}_h^T J_k=\mathbf{e}_k^T$. Instead, for all n the sum of the first and of the nth row of $A+JB$ determines $A+JB$, i.e. there exists a basis $\{J_k\}$ of γ with the property $(\mathbf{e}_1+\mathbf{e}_n)^T J_k=\mathbf{e}_k^T$.

The matrices of γ can be simultaneously diagonalized by a fast discrete transform. More precisely, for any value of n the following equality holds:

$$\begin{aligned} &\gamma = \{Gd(\mathbf{z})G^{-1} : \mathbf{z}\in\mathbb{C}^n\}, \\ &G_{ij} = \tfrac{1}{\sqrt{n}}\Big(\cos\tfrac{(2i-1)(2j-1)\pi}{2n} + \sin\tfrac{(2i-1)(2j-1)\pi}{2n}\Big),\ i,j=1,\ldots,n \end{aligned}$$

(see []). Note that the matrix G is real, symmetric, persymmetric and unitary. The fact that the matrix-vector product $G\mathbf{z}$ can be computed in at most $O(n\log_2 n)$ arithmetic operations follows from the representation of G_n, $G_n := G$, stated in Exercise G.

Note that for $n = 6$ the matrix G has ten zeros among its entries, and that these zeros are positioned as follows:

$$G = \begin{bmatrix} & & & & 0 & \\ & 0 & & 0 & & 0 \\ & & & & 0 & \\ & 0 & & & & \\ 0 & & 0 & & 0 & \\ & 0 & & & & \end{bmatrix}.$$

Thus each of the vectors $G^T \mathbf{e}_k$, $k = 1, \dots, 6$, has at least one zero entry, i.e. the matrix $d(G^T \mathbf{e}_k)^{-1}$ is never well defined. The latter assertion is yet true whenever $n = 2 + 4s$ $s \in \mathbb{Z}$. In other words, γ can be represented as

$$\gamma = \{G d(G^T \mathbf{z}) d(G^T \mathbf{e}_k)^{-1} G^{-1} : \mathbf{z} \in \mathbb{C}^n\}.$$

if and only $n \neq 2 + 4s$ $s \in \mathbb{Z}$ (for such n one can choose $k = 1$).

However, as one may guess from the above discussion (detailed, for $n = 6$), it can be easily shown that $[G^T(\mathbf{e}_1 + \mathbf{e}_n)]_k \neq 0$ $\forall\, k$ and $\forall\, n$. So, we have the following representation for γ

$$\gamma = \{G d(G^T \mathbf{z}) d(G^T(\mathbf{e}_1 + \mathbf{e}_n))^{-1} G^{-1} : \mathbf{z} \in \mathbb{C}^n\}$$

valid for all n (also for $n = 2 + 4s$ where $d(G^T \mathbf{e}_k)$ are not invertible). The latter formula confirms the fact that the matrices of γ are uniquely defined by the sum of their 1st and nth rows; in particular, since the sum of the first and of the nth row of the matrix $G d(G^T \mathbf{a}) d(G^T(\mathbf{e}_1 + \mathbf{e}_n))^{-1} G^{-1}$ is equal to $\mathbf{a}^T$, we can say that such matrix is exactly the matrix $\gamma(\mathbf{a})$ (already defined above for $n = 6$), i.e. the matrix of γ whose $(\mathbf{e}_1 + \mathbf{e}_n)$-row is $\mathbf{a}^T$.

It is now natural to introduce a class of spaces which include (besides the 1-spaces, like the Hessenberg algebras) also spaces like the algebra γ.

Definition. A subset $\mathcal{L}$ of $\mathbb{C}^{n \times n}$ is said to be a space in $\mathbb{V}$, if $\mathcal{L} = \text{Span}\,\{J_1, \dots, J_n\}$ with $\mathbf{v}^T J_k = \mathbf{e}_k^T$ for some $\mathbf{v} \in \mathbb{C}^n$. Given $\mathbf{z} \in \mathbb{C}^n$, the matrix $\sum_k z_k J_k \in \mathcal{L}$ is denoted by $\mathcal{L}(\mathbf{z})$. Since $\mathbf{v}^T \mathcal{L}(\mathbf{z}) = \mathbf{z}^T$, $\mathcal{L}(\mathbf{z})$ is called the matrix of $\mathcal{L}$ whose $\mathbf{v}$-row is $\mathbf{z}^T$.

Example. $\mathcal{L} = sd\, M = \{M d(\mathbf{z}) M^{-1} : \mathbf{z} \in \mathbb{C}^n\}$ is in $\mathbb{V}$ since

$$\mathcal{L} = \text{Span}\,\{J_1, \dots, J_n\},\ J_k = M d(M^T \mathbf{e}_k) d(M^T \mathbf{v})^{-1} M^{-1},$$

for any vector $\mathbf{v}$ such that $[M^T \mathbf{v}]_i \neq 0$ $\forall\, i$, and

$$\mathbf{v}^T J_k = \mathbf{v}^T M d(M^T \mathbf{e}_k) d(M^T \mathbf{v})^{-1} M^{-1} = \mathbf{e}_k^T M d(M^T \mathbf{v}) d(M^T \mathbf{v})^{-1} M^{-1} = \mathbf{e}_k^T.$$

Note that $\mathcal{L}(\mathbf{z}) = M d(M^T \mathbf{z}) d(M^T \mathbf{v})^{-1} M^{-1}$.

A matrix $X \in \mathbb{C}^{n \times n}$ is said to be non derogatory if the condition $p(X) = 0$, p polynomial, implies $\partial p \geq n$. Note that by the Cailey-Hamilton theorem the characteristic polynomial of X is null in X. So, X is non derogatory if and only if the set $\{p(X)\}$ of all polynomials in X has dimension n. In [] it is stated the following result, which proves that $\mathbb{V}$ is a wide class of spaces of matrices.

Theorem ND. Let X be a $n \times n$ matrix with complex entries. Then X is non derogatory if and only if $\{p(X)\} := \{p(X) : p \text{ polynomials}\}$ is in $\mathbb{V}$.

The Proposition here below collects several properties of the spaces in $\mathbb{V}$. They will be used (in the next section) in order to prove important properties of the best least squares fit in $\mathcal{L}$ of a matrix A, holding for all spaces $\mathcal{L}$ of a particular subclass of $\mathbb{V}$.

Proposition $\mathbb{V}$ *(properties of spaces in* $\mathbb{V}$*).* Let $\mathcal{L}$ be a space in $\mathbb{V}$, i.e. $\mathcal{L} = \text{Span}\,\{J_1, \dots, J_n\}$ with $\mathbf{v}^T J_k = \mathbf{e}_k^T$ for some $\mathbf{v} \in \mathbb{C}^n$.

(1) If $X \in \mathcal{L}$ and $\mathbf{v}^T X = \mathbf{0}^T$, then $X = 0$, thus $\mathbf{v}^T X = \mathbf{v}^T Y$, $X, Y \in \mathcal{L}$, implies $X = Y$

proof: $\mathbf{0}^T = \mathbf{v}^T X = \mathbf{v}^T \sum_k \alpha_k J_k = \sum_k \alpha_k \mathbf{e}_k^T = [\alpha_1 \ \cdots \ \alpha_n] \Rightarrow \alpha_k = 0 \ \forall\, k$.

(2) If $J_i X \in \mathcal{L}$, $X \in \mathbb{C}^{n\times n}$, then $J_i X = \sum_k [X]_{ik} J_k$

proof: there exist α_k such that $J_i X = \sum_k \alpha_k J_k$; multiplying the latter identity by $\mathbf{v}^T$ we have

$$\mathbf{e}_i^T X = \mathbf{v}^T J_i X = \sum_k \alpha_k \mathbf{e}_k^T = [\alpha_1 \ \cdots \ \alpha_n]$$

which implies $\alpha_k = [X]_{ik}$.

(3) Let $P_k \in \mathbb{C}^n$ be defined by $\mathbf{e}_s^T P_k = \mathbf{e}_k^T J_s$ (note that $\mathbf{e}_k^T = \mathbf{v}^T J_k = \sum_i v_i \mathbf{e}_i^T J_k = \sum_i v_i \mathbf{e}_k^T P_i = \mathbf{e}_k^T \sum_i v_i P_i$, and thus $\sum v_k P_k = I$). Then the following assertions are equivalent:

(i) $\mathcal{L}$ is closed under matrix multiplication

(ii) $J_i J_j = \sum_k [J_j]_{ik} J_k \ \forall\, i, j$

(iii) $P_r J_j = J_j P_r \ \forall\, r, j$

(iv) $P_k P_r = \sum_i [P_r]_{ki} P_i$

proof: The implication (i) $\Rightarrow$ (ii) follows from (2) for $X = J_j$. The opposite implication is obvious. The fact that conditions (ii) and (iii) are equivalent follows by taking the (r, s) entry of the equality in (ii):

$$\mathbf{e}_i^T P_r J_j \mathbf{e}_s = \mathbf{e}_r^T J_i J_j \mathbf{e}_s = \sum_k [J_j]_{ik} [J_k]_{rs} = \sum_k [J_j]_{ik} [P_r]_{ks} = [J_j P_r]_{is}.$$

The fact that conditions (iii) and (iv) are equivalent follows from the identities:

$$[P_k P_r]_{ms} = [J_m P_r]_{ks} = [P_r J_m]_{ks} = [J_k J_m]_{rs} = \sum_i [J_k]_{ri} [J_m]_{is} = \sum_i [P_r]_{ki} [P_i]_{ms}.$$

(3.5) If $I \in \mathcal{L}$, then $\sum_i v_i J_i = I$ and $\mathbf{v}^T P_k = \mathbf{e}_k^T$, i.e. also the space $\text{Span}\,\{P_1, \dots, P_n\}$ is in $\mathbb{V}$ (with the same $\mathbf{v}$)

proof: both I and $\sum_i v_i J_i$ have $\mathbf{v}^T$ as $\mathbf{v}$-row, and both, by assumption, are in $\mathcal{L}$, so they must be equal; moreover, we have

$$\mathbf{v}^T P_k = \sum_i v_i \mathbf{e}_i^T P_k = \sum_i v_i \mathbf{e}_k^T J_i = \mathbf{e}_k^T \sum_i v_i J_i = \mathbf{e}_k^T.$$

(4) If $\mathcal{L}$ is closed under matrix multiplication, then

$$\mathcal{L}(\mathcal{L}(\mathbf{z})^T \mathbf{z}') = \mathcal{L}(\mathbf{z}')\mathcal{L}(\mathbf{z}), \ \forall\, \mathbf{z}, \mathbf{z}' \in \mathbb{C}^n$$

proof: since $\mathcal{L}$ is closed, the matrix $\mathcal{L}(\mathbf{z}')\mathcal{L}(\mathbf{z})$ is in $\mathcal{L}$; moreover, its $\mathbf{v}$-row is ${\mathbf{z}'}^T\mathcal{L}(\mathbf{z})$; the thesis follows from the fact that also $\mathcal{L}(\mathcal{L}(\mathbf{z})^T\mathbf{z}')$ is the matrix of $\mathcal{L}$ whose $\mathbf{v}$-row is ${\mathbf{z}'}^T\mathcal{L}(\mathbf{z})$.

(5) Assume $I \in \mathcal{L}$ and $\mathcal{L}$ closed under matrix multiplication. Then $X \in \mathcal{L}$ is non singular if and only if $\exists\, \mathbf{z} \in \mathbb{C}^n$ such that $\mathbf{z}^T X = \mathbf{v}^T$; in this case $X^{-1} = \mathcal{L}(\mathbf{z})$

proof: by inspecting the kth row of the matrix $\mathcal{L}(\mathbf{z})X$, and applying properties (3) and (3.5), we obtain the identities

$$\mathbf{e}_k^T\mathcal{L}(\mathbf{z})X = \mathbf{e}_k^T\sum_s z_s J_s X = \sum_s z_s \mathbf{e}_s^T P_k X = \mathbf{z}^T P_k X = \mathbf{z}^T X P_k = \mathbf{v}^T P_k = \mathbf{e}_k^T,$$

or, equivalently, the equality $\mathcal{L}(\mathbf{z})X = I$, which implies that X is non singular and $X^{-1} = \mathcal{L}(\mathbf{z})$.

The best least squares fit to A in $\mathcal{L} \subset \mathbb{C}^{n\times n}$

Given a subspace $\mathcal{L} \subset \mathbb{C}^{n\times n}$ and a $n\times n$ matrix A, it is well defined $\mathcal{L}_A$, the projection on $\mathcal{L}$ of A. In the following theorem we state some assumptions on $\mathcal{L}$ (in particular we consider n-dimensional subspaces of $\mathbb{C}^{n\times n}$) which assure that $\mathcal{L}_A$ is hermitian whenever A is, and that the eigenvalues of $\mathcal{L}_A$ are bounded by those of A. Such assumptions imply that $\mathcal{L} \in \mathbb{V}$.

Theorem $\mathcal{L}_A$. Assumptions:

$\mathcal{L} \subset \mathbb{C}^{n\times n}$, $I \in \mathcal{L}$, $\mathcal{L} =$ Span$\{J_1,\dots,J_n\}$ with J_k such that

$$J_i^H J_j = \sum_{k=1}^n \overline{[J_k]_{ij}} J_k,\ \ i,j = 1,\dots,n. \qquad (*)$$

$A \in \mathbb{C}^{n\times n}$.

$\mathcal{L}_A \in \mathcal{L}$, $\|A - \mathcal{L}_A\|_F \le \|A - X\|_F$, $\forall\, X \in \mathcal{L}$ (such matrix $\mathcal{L}_A$ is well defined since $\mathbb{C}^{n\times n}$ is a Hilbert space with respect the norm $\|\cdot\|_F$ induced by the inner product $(A,B)_F = \sum_{ij}\overline{a_{ij}}b_{ij}$ and $\mathcal{L}$ is a subspace of $\mathbb{C}^{n\times n}$).

Thesis: If $A = A^H$, then $\mathcal{L}_A = \mathcal{L}_A^H$ and $\min\lambda(A) \le \lambda(\mathcal{L}_A) \le \max\lambda(A)$.

Note: if A is real symmetric, then $\mathcal{L}_A$ is in general hermitian; it is real symmetric under the further condition that $\mathcal{L}$ is spanned by real matrices (prove it!).

Note: we shall see that the hypotheses of Theorem $\mathcal{L}_A$ are satisfied by spaces of the type $\{Md(\mathbf{z})M^{-1} : \mathbf{z} \in \mathbb{C}^n\}$ if $M^H M$ is diagonal and its diagonal entries are positive; however, the same hypotheses can be satisfied also by non commutative spaces (we shall see an example, for others see []), so also in the latter cases we can say that the conclusions of Theorem $\mathcal{L}_A$ hold.

Applications of Theorem $\mathcal{L}_A$. If the conditions of Theorem $\mathcal{L}_A$ are satisfied, then $\mathcal{L}_A$ is positive definite (i.e. $\mathcal{L}_A = \mathcal{L}_A^H$ and $\mathbf{z} \in \mathbb{C}^n$ $\mathbf{z} \ne \mathbf{0} \Rightarrow \mathbf{z}^H\mathcal{L}_A\mathbf{z}$ positive) whenever A is positive definite. So, in order to solve the linear system

$$A\mathbf{x} = \mathbf{b},\ \ A \text{ positive definite}$$

we can solve the equivalent system

$$\mathcal{L}_A^{-1}A\mathbf{x} = \mathcal{L}_A^{-1}\mathbf{b}$$

whose coefficient matrix has real and positive eigenvalues, often better distributed than those of A (this results, for example when solving $A\mathbf{x} = \mathbf{b}$ iteratively, in less iterations).

Moreover, if the conditions of Theorem $\mathcal{L}_A$ are satisfied and $\mathcal{L}$ is spanned by real matrices, then then $\mathcal{L}_A$ is real positive definite (i.e. $\mathcal{L}_A = \mathcal{L}_A^T$, $\mathcal{L}_A \in \mathbb{R}^{n\times n}$, and $\mathbf{z} \in \mathbb{C}^n$ $\mathbf{z} \neq \mathbf{0} \Rightarrow \mathbf{z}^H\mathcal{L}_A\mathbf{z}$ positive) whenever A is real positive definite. That is, the following implications hold

$$\begin{aligned}
&B_k \text{ real positive definite } \Rightarrow \\
&\mathcal{L}_{B_k} \text{ real positive definite } \Rightarrow \\
&\varphi(\mathcal{L}_{B_k}, \mathbf{s}_k, \mathbf{y}_k) \text{ real positive definite} \Rightarrow \\
&\mathcal{L}_{\varphi(\mathcal{L}_{B_k}, \mathbf{s}_k, \mathbf{y}_k)} \text{ real positive definite}
\end{aligned}$$

(provided $\mathbf{s}_k^T\mathbf{y}_k$ is positive), and thus both the S and NS LQN search directions,

$$\mathbf{d}_{k+1} = -\varphi(\mathcal{L}_{B_k}, \mathbf{s}_k, \mathbf{y}_k)^{-1}\nabla f(\mathbf{x}_{k+1}) \quad \text{and} \quad \mathbf{d}_{k+1} = -\mathcal{L}^{-1}_{\varphi(\mathcal{L}_{B_k}, \mathbf{s}_k, \mathbf{y}_k)}\nabla f(\mathbf{x}_{k+1}),$$

are well defined descent directions in $\mathbf{x}_{k+1}$ for the function $f : \mathbb{R}^n \to \mathbb{R}$ (see [] for the definitions of B_k, $\mathbf{s}_k$, $\mathbf{y}_k$, φ, S and NS LQN).

proof (of Theorem $\mathcal{L}_A$): The matrix $\mathcal{L}_A = \sum_s \alpha_s J_s$ is uniquely defined by the following condition

$$(X, A - \mathcal{L}_A)_F = 0, \ X \in \mathcal{L}$$

or, equivalently, by the n conditions $(J_k, A - \sum_s \alpha_s J_s)_F = 0$, $k = 1, \dots, n$, which can be rewritten as follows

$$\sum_{s=1}^{n}(J_k, J_s)_F\alpha_s = (J_k, A)_F, \ k = 1, \dots, n.$$

In other words we have the formula

$$\mathcal{L}_A = \sum_s [B^{-1}\mathbf{c}]_s J_s, \ B_{ks} = (J_k, J_s)_F, \ c_k = (J_k, A)_F, \ k, s = 1, \dots, n.$$

Remark. B is positive definite, i.e. $B = B^*$ and $\mathbf{z}^H B\mathbf{z} > 0 \ \forall \mathbf{z} \in \mathbb{C}^n$ $\mathbf{z} \neq \mathbf{0}$.

proof: $\overline{B_{ks}} = \overline{(J_k, J_s)_F} = (J_s, J_k)_F = B_{sk}$, that is, B is a hermitian matrix. Moreover, since $0 < (\sum_s z_s J_s, \sum_s z_s J_s)_F = \sum_{k,s} \overline{z_k} z_s (J_k, J_s)_F = \mathbf{z}^H B\mathbf{z}$ whenever $\mathbf{z} \neq \mathbf{0}$, the matrix B is also positive definite.

Remark. Let $v_k \in \mathbb{C}$ be such that $I = \sum_k v_k J_k$ (such v_k exist because $I \in \mathcal{L}$). Then the vector $\mathbf{v}$ whose entries are the v_k satisfies the equalities $\mathbf{v}^T J_k = \mathbf{e}_k^T$, thus $\mathcal{L} \in \mathbb{V}$ and all results stated for spaces in $\mathbb{V}$ hold for our space $\mathcal{L}$.

proof: Multiply (*) by $\overline{v_i}$ and sum on i:

$$\overline{v_i}J_i^H J_j = \sum_k \overline{v_i}\overline{[J_k]_{ij}}J_k, \ J_j = \sum_k (\sum_i \overline{v_i}\overline{[J_k]_{ij}})J_k = \sum_k (\mathbf{v}^H\overline{J_k}\mathbf{e}_j)J_k.$$

This implies $\mathbf{v}^H\overline{J_k}\mathbf{e}_j = 0$ if $k \neq j$ and $\mathbf{v}^H\overline{J_k}\mathbf{e}_j = 1$ if $k = j$, i.e. $\mathbf{v}^T J_k = \mathbf{e}_k^T$.

As an immediate consequence of the above two Remarks, we have that $\mathcal{L}_A$ is the matrix of $\mathcal{L}$ whose $\mathbf{v}$-row is $(B^{-1}\mathbf{c})^T$,

$$\mathcal{L}_A = \sum_s [B^{-1}\mathbf{c}]_s J_s = \mathcal{L}(B^{-1}\mathbf{c}),$$

and, moreover,

$$\mathcal{L}_A = \mathcal{L}(B^{-1}\mathbf{c}) = \mathcal{L}((B^H)^{-1}\mathbf{c}) = \mathcal{L}((\overline{B}^{-1})^T\mathbf{c}).$$

Remark. $\mathcal{L}$ is closed under conjugate transposition.

proof: multiply (*) by v_j and sum on j:

$$J_i^H v_j J_j = \sum_k v_j \overline{[J_k]_{ij}} J_k,\ \ J_i^H = \sum_k (\sum_j v_j \overline{[J_k]_{ij}}) J_k \ \Rightarrow\ J_i^H \in \mathcal{L}.$$

The latter Remark yields part of the thesis of Theorem $\mathcal{L}_A$, because it implies that $\mathcal{L}_A^H \in \mathcal{L}$, and this fact together with the equalities

$$\|A - \mathcal{L}_A\|_F = \|A^H - \mathcal{L}_A^H\|_F = \|A - \mathcal{L}_A^H\|_F$$

(remember that our A is hermitian!) and the unicity of the best approximation of A, yield the identity $\mathcal{L}_A = \mathcal{L}_A^H$. In other words, under our conditions on $\mathcal{L}$ the projection on $\mathcal{L}$ of a hermitian matrix is hermitian too.

Remark. $\mathcal{L}$ is closed under matrix multiplication ($\mathcal{L}$ is a matrix algebra).

proof: the set $\{J_i^H\}$ forms an alternative basis for $\mathcal{L}$ (prove it!), thus there exist $z_i^{(s)} \in \mathbb{C}$ such that $J_s = \sum_i z_i^{(s)} J_i^H$. Multiply (*) by $z_i^{(s)}$ and sum on i,

$$z_i^{(s)} J_i^H J_j = \sum_k z_i^{(s)} \overline{[J_k]_{ij}} J_k,\ \ J_s J_j = \sum_k (\sum_i z_i^{(s)} \overline{[J_k]_{ij}}) J_k,$$

to observe that $J_s J_j \in \mathcal{L}$.

Remark. $\overline{B} = \sum_k \overline{\mathrm{tr}\,(J_k)} J_k$, thus $\overline{B} \in \mathcal{L}$, and, since $\overline{B}$ is non singular (it is positive definite!), by the result $\mathbb{V}$ (5) also the matrix $\overline{B}^{-1}$ is in $\mathcal{L}$.

proof: by equality (*) we have:

$$\begin{array}{rcl} B_{ij} & = & (J_i, J_j)_F = \sum_{r,t} \overline{[J_i]_{rt}} [J_j]_{rt} = \sum_{r,t} [J_i^H]_{tr} [J_j]_{rt} \\ & = & \sum_t [J_i^H J_j]_{tt} = \sum_t \sum_k \overline{[J_k]_{ij}} [J_k]_{tt} = \sum_k \mathrm{tr}\,(J_k) \overline{[J_k]_{ij}}. \end{array}$$

The latter two Remarks, together with $\mathbb{V}$ (4), let us rewrite again $\mathcal{L}_A$ as follows

$$\mathcal{L}_A = \ldots = \mathcal{L}((\overline{B}^{-1})^T\mathbf{c}) = \mathcal{L}(\mathbf{c})\overline{B}^{-1}.$$

Now note that there exists a hermitian matrix M such that $M^2 = \overline{B}^{-1}$, and that the matrices $\mathcal{L}_A$ and $M\mathcal{L}(\mathbf{c})M$ have the same eigenvalues (by the last representation of $\mathcal{L}_A$ they are similar!). So, if $\lambda(\mathcal{L}_A)$ is the generic eigenvalue of $\mathcal{L}_A$, then there exists $\mathbf{x} \in \mathbb{C}^n$ $\|\mathbf{x}\|_2 = 1$ such that

$$\lambda(\mathcal{L}_A) = \mathbf{x}^H M \mathcal{L}(\mathbf{c}) M \mathbf{x} = (M\mathbf{x})^H \mathcal{L}(\mathbf{c})(M\mathbf{x}).$$

Remark. If $\mathbf{z} \in \mathbb{C}^n$, then $\mathbf{z}^H \mathcal{L}(\mathbf{c})\mathbf{z} = \sum_k (P_k^H \mathbf{z})^H A (P_k^H \mathbf{z})$.

proof: here again equality (*) is fundamental:

$$\begin{array}{rcl} \mathbf{z}^H \mathcal{L}(\mathbf{c})\mathbf{z} & = & \mathbf{z}^H (\sum_k (J_k, A)_F J_k)\mathbf{z} \\ & = & \sum_{i,j=1}^n \overline{z_i} z_j \sum_{r,t=1}^n a_{rt} \overline{\sum_k [J_k]_{ij} \overline{[J_k]_{rt}}} \\ & = & \sum_{i,j=1}^n \overline{z_i} z_j \sum_{r,t=1}^n a_{rt} \overline{[J_i^H J_j]_{rt}} \\ & = & \sum_{r,t=1}^n a_{rt} \overline{\sum_{i,j=1}^n \overline{z_i} z_j \sum_k [J_i^H]_{rk} [J_j]_{kt}} \\ & = & \sum_k \sum_{r,t} a_{rt} \overline{(\sum_i z_i [P_k^H]_{ri})} (\sum_j z_j [P_k^H]_{tj}) \\ & = & \sum_k \sum_{r,t} a_{rt} \overline{(P_k^H \mathbf{z})_r} (P_k^H \mathbf{z})_t. \end{array}$$

By the above Remark we have:

$$\begin{array}{rcl}\lambda(\mathcal{L}_A) & = & \sum_k (P_k^H M\mathbf{x})^H A(P_k^H M\mathbf{x}) \le \max\lambda(A)\sum_k (P_k^H M\mathbf{x})^H (P_k^H M\mathbf{x}) \\ & = & \max\lambda(A)\mathbf{x}^H M(\sum_k P_k P_k^H)M\mathbf{x}.\end{array}$$

But the matrix in the brackets is nothing else the matrix $\overline{B}$:

Remark. $\overline{B} = \sum_k P_k P_k^H$

proof:

$$B_{ij} = (J_i, J_j)_F = \sum_{r,s} \overline{[J_i]_{rs}}[J_j]_{rs} = \sum_{r,s} \overline{[P_r]_{is}}[P_r]_{js} = \sum_{r,s} \overline{[P_r]_{is}}[P_r^T]_{sj} = \sum_r [\overline{P_r}P_r^T]_{ij}.$$

We can now conclude one of the inequalities (stated in Theorem $\mathcal{L}_A$) satisfied by the eigenvalues of A and $\mathcal{L}_A$:

$$\lambda(\mathcal{L}_A) \le \max\lambda(A)\mathbf{x}^H M M^{-2} M\mathbf{x} = \max\lambda(A)\mathbf{x}^H\mathbf{x} = \max\lambda(A).$$

Analogously, one can prove that $\lambda(\mathcal{L}_A) \ge \min\lambda(A)$. $\square$

Exercise DG. Prove that the space

$$\mathcal{L} = \{\begin{bmatrix} X & JY \\ JY & X \end{bmatrix} : \ X, Y \ \frac{n}{2} \times \frac{n}{2} \text{ circulants}\}$$

satisfies the hypothesis of Theorem $\mathcal{L}_A$, i.e. $I \in \mathcal{L}$, $\mathcal{L} = \text{Span}\{J_1, \dots, J_n\}$ with J_k linearly independent such that

$$J_i^H J_j = \sum_{k=1}^n \overline{[J_k]_{ij}} J_k, \ i, j = 1, \dots, n.$$

Thus, the projection $\mathcal{L}_A$ on $\mathcal{L}$ is hermitian and such that $\min\lambda(A) \le \lambda(\mathcal{L}_A) \le \max\lambda(A)$ whenever A is hermitian. Note that $\mathcal{L}$ is not commutative.

Proposition c$\mathbb{V}$ *(properties of commutative spaces in* $\mathbb{V}$*)* [mitia]. Let $\mathcal{L}$ be a space in $\mathbb{V}$, i.e. $\mathcal{L} = \text{Span}\{J_1, \dots, J_n\}$ with $\mathbf{v}^T J_k = \mathbf{e}_k^T$ for some $\mathbf{v} \in \mathbb{C}^n$. Assume that $\mathcal{L}$ is commutative. Then

(1) $\mathbf{e}_i^T J_j = \mathbf{e}_j^T J_i$, $\forall\, i, j$, and thus $J_k = P_k$

proof: $J_i J_j = J_j J_i \Rightarrow \mathbf{v}^T J_i J_j = \mathbf{v}^T J_j J_i$, and the definition $\mathbf{v}^T J_k = \mathbf{e}_k^T$ yields the thesis.

(2) $\mathbf{z}^T \mathcal{L}(\mathbf{z}') = \mathbf{z}'^T \mathcal{L}(\mathbf{z})$

proof: by (1) we have

$$\mathbf{z}^T \mathcal{L}(\mathbf{z}') = \sum_i z_i \mathbf{e}_i^T \sum_k z'_k J_k = \sum_i z_i \sum_k z'_k \mathbf{e}_k^T J_i = \sum_k z'_k \mathbf{e}_k^T \sum_i z_i J_i.$$

(3) $I = \mathcal{L}(\mathbf{v}) \in \mathcal{L}$

proof: note that $\mathcal{L}(\mathbf{v}) = \sum v_i J_i \in \mathcal{L}$ and $\mathbf{e}_k^T \mathcal{L}(\mathbf{v}) = \sum_i v_i \mathbf{e}_i^T J_k = \mathbf{v}^T J_k = \mathbf{e}_k^T$, so $I = \mathcal{L}(\mathbf{v}) \in \mathcal{L}$.

(4) $\mathcal{L}$ is closed under matrix multiplication

proof: from (1) we have that $J_k = P_k$, thus $P_k J_s = J_k J_s = J_s J_k = J_s P_k$, which is one of the necessary and sufficient conditions for the multiplicative closure.

Example of commutative $\mathcal{L} \in \mathbb{V}$. Let M be a non singular $n \times n$ matrix with complex entries, and set $\mathcal{L} = sd\,M = \{Md(\mathbf{z})M^{-1} : \mathbf{z} \in \mathbb{C}^n\}$. Note that $\mathcal{L} \in \mathbb{V}$, in fact if $\mathbf{v}$ is any vector such that $[M^T\mathbf{v}]_k \neq 0$, $\forall\, k$, then the matrices $J_k = Md(M^T\mathbf{e}_k)d(M^T\mathbf{v})^{-1}M^{-1}$ satisfy the identities $\mathbf{v}^T J_k = \mathbf{e}_k^T$ and span $\mathcal{L}$. We have, for $\mathcal{L}$, the following alternative representation:

$$\mathcal{L} = \{Md(M^T\mathbf{z})d(M^T\mathbf{v})^{-1}M^{-1} : \mathbf{z} \in \mathbb{C}^n\}.$$

It is clear that the matrix of $\mathcal{L}$ whose $\mathbf{v}$-row is $\mathbf{z}^T$ is

$$\mathcal{L}(\mathbf{z}) = Md(M^T\mathbf{z})d(M^T\mathbf{v})^{-1}M^{-1}.$$

Obviously, $\mathcal{L}$ is commutative.

Proposition ch$\mathbb{V}$ *(properties of commutative, closed under conjugate transposition spaces in* $\mathbb{V}$*)* [stefano]. Let $\mathcal{L}$ be a space in $\mathbb{V}$, i.e. $\mathcal{L} = \text{Span}\,\{J_1, \ldots, J_n\}$ with $\mathbf{v}^T J_k = \mathbf{e}_k^T$ for some $\mathbf{v} \in \mathbb{C}^n$. Assume that $\mathcal{L}$ is commutative and closed under conjugate transposition. Then, besides the above c$\mathbb{V}$ (1),(2),(3),(4), we have

$$J_i^H J_j = \sum_k \overline{[J_k]_{ij}} J_k, \ \ i, j = 1, \ldots, n$$

proof: since $J_i^H \in \mathcal{L}$ and $\mathcal{L}$ is commutative, one has $J_i^H J_j = J_j J_i^H$; since $\mathcal{L}$ is closed under matrix multiplication and $J_i^H \in \mathcal{L}$, one has that $J_j J_i^H \in \mathcal{L}$; by $\mathbb{V}$ (2), it follows that

$$J_i^H J_j = J_j J_i^H = \sum_k [J_i^H]_{jk} J_k = \sum_k \overline{[J_i]_{kj}} J_k = \sum_k \overline{[J_k]_{ij}} J_k,$$

where in the latter identity we have used property c$\mathbb{V}$(1).

Example of commutative, closed under conjugate transposition $\mathcal{L} \in \mathbb{V}$. Let M be a non singular $n \times n$ matrix with complex entries, and set $\mathcal{L} = sd\,M = \{Md(\mathbf{z})M^{-1} : \mathbf{z} \in \mathbb{C}^n\}$. We already know that $\mathcal{L}$ is a space in $\mathbb{V}$ which is commutative. We want to prove that $\mathcal{L}$ *is closed under conjugate transposition if and only if* $M^H M$ *is diagonal with positive diagonal entries.*

One implication is easy: assume $M^H M = D$, D_{ii} positive $\forall\, i$, $D_{ij} = 0$ $i \neq j$; then

$$(Md(\mathbf{z})M^{-1})^H = (M^H)^{-1}d(\overline{\mathbf{z}})M^H = MD^{-1}d(\overline{\mathbf{z}})DM^{-1} = Md(\overline{\mathbf{z}})M^{-1} \ \in \ \mathcal{L}.$$

Now assume that $\mathcal{L}$ is closed under conjugate transposition. Thus, for any $\mathbf{z} \in \mathbb{C}^n$ there exists $\mathbf{w} \in \mathbb{C}^n$ such that $(Md(\mathbf{z})M^{-1})^H = Md(\mathbf{w})M^{-1}$. But this implies

$$d(\overline{\mathbf{z}})C = Cd(\mathbf{w}), \ C = M^H M,$$

or, equivalently, $c_{ij}\overline{z}_i = c_{ij}w_j$. Assume the z_i distinct. Then the identities $c_{i1}\overline{z}_i = c_{i1}w_1$, $i = 1, \ldots, n$, imply $c_{i1} = 0$ for all i except one of them, say i_1 (otherwise, $c_{i1} \neq 0$ and $c_{k1} \neq 0$, $i \neq k$, would imply $w_1 = \overline{z}_i = \overline{z}_k$!). Analogously, the identities $c_{i2}\overline{z}_i = c_{i2}w_2$, $i = 1, \ldots, n$, imply $c_{i2} = 0$ for all i except one of them, say i_2, and such i_2 must be different fom i_1 otherwise C would be singular. Proceeding in this way one concludes that

$$C = DR, \ D \text{ diagonal}, \ R \text{ permutation}.$$

The fact that C is hermitian implies that the permutation matrix R must be symmetric. The fact that the diagonal entries of C cannot be zero implies that R is the identity. Finally, the fact that C is positive definite implies that D_{ii} are real and positive.

Question. Find an example of $\mathcal{L}$ satisfying the assumptions of Proposition ch$\mathbb{V}$ which is not of the type $\{Md(\mathbf{z})M^{-1} : \mathbf{z} \in \mathbb{C}^n\}$, $M^H M$ diagonal with positive diagonal entries.

Exercise. Let $\mathcal{L}$ be in $\mathbb{V}$ and closed under matrix multiplication. Prove that
(i) If the matrices in $\mathcal{L}$ are symmetric, then $\mathcal{L}$ is commutative
(ii) If the matrices in $\mathcal{L}$ are persymmetric, then $\mathcal{L}$ is commutative

proof: by the assumptions on $\mathcal{L}$, we have, in the symmetric case,

$$J_s J_k = J_s^T J_k^T = (J_k J_s)^T = J_k J_s,$$

and, in the persymmetric case,

$$J_s J_k = J_s J J J_k = J J_s^T J_k^T J = J((J_k J_s)^T)J = J(J J_k J_s J)J = J_k J_s.$$

EXERCISE.

$$P_\xi = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ \xi & & & & 0 \end{bmatrix}, \ \xi \neq 0$$

(1) If $\rho^n = \xi$ and $\omega^n = 1$, then

$$P_\xi \begin{bmatrix} 1 \\ \rho\omega^j \\ \\ \rho^{n-1}\omega^{(n-1)j} \end{bmatrix} = \rho\omega^j \begin{bmatrix} 1 \\ \rho\omega^j \\ \\ \rho^{n-1}\omega^{(n-1)j} \end{bmatrix}, \ j = 0, 1, \dots, n-1.$$

Thus, if $W = (\omega^{ij})_{i,j=0}^{n-1}$ and $D_{1\rho^{n-1}} = \operatorname{diag}(\rho^i,\, i = 0, \dots, n-1)$, then

$$P_\xi(D_{1\rho^{n-1}}W) = (D_{1\rho^{n-1}}W)\rho D_{1\omega^{n-1}}, \ D_{1\omega^{n-1}} = \operatorname{diag}(\omega^j,\, j = 0, \dots, n-1).$$

(2) If, moreover, $|\xi| = 1$ and $\omega^i \neq 1$ $0 < i < n$, then $U = \frac{1}{\sqrt{n}} D_{1\rho^{n-1}} W$ is unitary and

$$\begin{aligned} C_\xi &:= H_{P_\xi} = \{\textstyle\sum_{k=1}^n z_k P_\xi^{k-1} : z_k \in \mathbb{C}\} \\ &= \{Ud(\mathbf{z})U^* : \mathbf{z} \in \mathbb{C}^n\} = \{Ud(U^T\mathbf{z})d(U^T\mathbf{e}_1)^{-1}U^{-1} : \mathbf{z} \in \mathbb{C}^n\}. \end{aligned}$$

The matrix $C_\xi(\mathbf{a}) := \sum_{k=1}^n a_k P_\xi^{k-1} = Ud(U^T\mathbf{a})d(U^T\mathbf{e}_1)^{-1}U^{-1}$ is the ξ-circulant matrix whose first row is $\mathbf{a}^T$:

$$C_\xi(\mathbf{a}) = \begin{bmatrix} a_1 & a_2 & & a_{n-1} & a_n \\ \xi a_n & a_1 & & & a_{n-1} \\ & & & & \\ \xi a_3 & & & & a_2 \\ \xi a_2 & \xi a_3 & & \xi a_n & a_1 \end{bmatrix}.$$

EXERCISE.

Set $X = P_1 + P_1^T$.

1) Find a convenient representation for the set C^S of all polynomials in X, and deduce the dimension of C^S.

2) Prove that any matrix of the form $C_1 + JC_2$, C_1, C_2 circulants, J anti-identity, belongs to the space $\{A \in \mathbb{C}^{n\times n} : AX = XA\}$. Deduce a lower bound for the dimension of $\{A \in \mathbb{C}^{n\times n} : AX = XA\}$.

Repeat the exercise for $X = P_{-1} + P_{-1}^T$.

EXERCISE

1) Write precisely an algorithm that computes the matrix-vector product $T \cdot \mathbf{z}$, $T = (t_{i-j})_{i,j=1}^n$, in $O(n \log_2 n)$ arithmetic operations.

2) Write an algorithm that computes the matrix-vector product $(T^{-1}) \cdot \mathbf{z}$ in $O(n \log_2 n)$ arithmetic operations (after preprocessing on T).

EXERCISE

Do exercise DG

EXERCISE Prove Theorem DD

EXERCISE Let T be a $n \times n$ Toeplitz matrix

1) Prove that $TP_0 - P_0T$ has rank at most 2

2) Let H be a Hankel matrix, i.e. $H = (h_{i+j-2})_{i,j=1}^n$, and show that $(T + H)(P_0 + P_0^T) - (P_0 + P_0^T)(T + H)$ has rank at most 4

3) Write the matrix of the algebra τ whose first row is $[2\ \ -1\ 0\ 0\ \cdots\ 0]$, and observe that it is a Toeplitz matrix. Call A such matrix and compute A^{-1} explicitly by using the fact that $A^{-1} \in \tau$ (so, it is sufficient to compute its first row). Is A^{-1} a Toeplitz matrix ?

4) Do exercise Ttau

EXERCISE Under the assumptions of Theorem $\mathcal{L}_A$

1) prove the last of the following identities

$$\mathcal{L}_A = \mathcal{L}(B^{-1}\mathbf{c}) = \mathcal{L}(\mathbf{c})\overline{B}^{-1} = \overline{B}^{-1}\mathcal{L}(\mathbf{c})$$

(hint: find an expression of $\overline{B}$ in terms of the matrices P_k, $\mathbf{e}_s^T P_k = \mathbf{e}_k^T J_s\ \forall\, s, k$)

2) prove that B is positive definite (besides hermitian)

EXERCISE (0) Let $\mathcal{L} \in \mathbb{C}^{n\times n}$, $\mathcal{L} =$ Span $\{J_1, \ldots, J_n\}$ with $\mathbf{v}^T J_k = \mathbf{e}_k^T$ for some vector $\mathbf{v}$ (that is, $\mathcal{L} \in \mathbb{V}$). Assume that $\mathcal{L}$ is closed under matrix multiplication, that $I \in \mathcal{L}$, and that $J_i^H = \alpha_i J_{t_i}$, $|\alpha_i| = 1\ \forall\, i$, for some $t_i \in \{1, 2, \ldots, n\}$.

(1) Prove that $J_i^H J_j = \sum_k \overline{[J_k]_{ij}} J_k$, $\forall\, i, j$.

(2) Assume that $\{1, 2, \ldots, n\}$ is a group with identity element 1. Prove that

$$\mathcal{L} = \{A \in \mathbb{C}^{n\times n} : a_{ij} = a_{1,i^{-1}j}\}$$

satisfies the assumptions (0).

(3) Prove that the space $\mathcal{L}$ spanned by the matrices $J_1 = I$,

$$J_2 = \begin{bmatrix} & 1 & & \\ 1 & & & \\ & & & \mathbf{i} \\ & & -\mathbf{i} & \end{bmatrix},\ J_3 = \begin{bmatrix} & & 1 & \\ & & & -\mathbf{i} \\ 1 & & & \\ & \mathbf{i} & & \end{bmatrix},\ J_4 = \begin{bmatrix} & & & 1 \\ & & \mathbf{i} & \\ & -\mathbf{i} & & \\ 1 & & & \end{bmatrix},$$

satisfies the assumptions (0). Prove that $\mathcal{L}$ is not commutative.

EXERCISE Let T be a symmetric Toeplitz matrix.

(1) Compute the first row of C_T where C is the space of circulant matrices

(2) Compute the first row of τ_T where τ is the space of tau matrices

(3) Compute the first row of $\mathcal{L}_T$ where $\mathcal{L}$ is the space in the point (3) of the previous exercise (so, $n = 4$)

(4) Compute the vector $(\mathbf{e}_1 + \mathbf{e}_6)^T \gamma_T$ where γ is the space of all 6×6 gamma matrices (i.e. assume $n = 6$).

REFERENCES

Linear Algebra Appl. 229 (1995), 49-99

Five rows of SIAM J. Matrix Anal. Appl. 21 (2000), 646-667

Linear Algebra Appl. 335 (2001), 1-54

For Hartley-type transforms and algebras (as G and γ) Linear Algebra Appl. 366 (2003), 65-85

For an intersection with prof. Ikramov research interests Computational Math. and Mathematical Physics 38 (1998), 1026-1035

Acknowledgements *To all the participants, Grazie! Prof*