Insights on some theoretical results related to the convergence of power and Gauss-Seidel iterations

While studying the rate of convergence of power method applied to A^{-1}, where A is a $n \times n$ not diagonalizable invertible matrix, one needs to consider the matrix Y which transforms A^{-1} into Jordan canonical form, and to express such Y in terms of the matrix X which transforms A into Jordan canonical form. The matrix Y turns out to be equal to XM with M block diagonal where the diagonal blocks have an interesting upper triangular Tartaglia-Toeplitz structure.

While searching for linear systems $Ax = b$ which are solvable via the Gauss-Seidel (G.-S.) method even if the standard sufficient conditions on A for G.-S. convergence are not satisfied, one meets an interesting class of 2×2-block matrices A, for which G.-S. converges for A every time it converges for its two diagonal blocks, of order r and $n-r$, and $|a_{r+1,1}a_{1,r+1}/a_{1,1}a_{r+1,r+1}| < 1$.

The above are two examples of how in-depth studies of classical subjects of numerical mathematics can lead to new interesting remarks.

E. Bozzo , P. Deidda, C. Di Fiore, The Jordan and Frobenius pairs of the inverse, Linear and Multilinear Algebra, 2023

I am Carmine Di Fiore, associate professor of Numerical Analysis in Rome “Tor Vergata” University.

I am happy to be here, in Shenzhen, China, at the CMC Faculty of the University MSU-BIT (Moscow State University - Beijing Institute of Technology), together with you, participants of the Scuola-Convegno “Tensor Methods in Mathematics and Data Science”, November 11-20, 2024.
I feel with us also my sister Maria Rita, Professor of Mathematics in school, thanks to whom I became professor in university.

I thank very much Professor Eugene Tyrtyshnikov, who invited me to this event.

I can say that I am not an expert in the field “Tensor Methods in Mathematics and Data Science”, in fact I should have attended the School, the first part of the event, as a student (when we will be back in Rome, I'll ask to Damiano to tell me some of the subjects he has learnt). However, I was happy to accept the invitation in order to meet again my two friends out of Italy, Eugene here in Shenzhen, and Professor Raymond Chan here and in Hong Kong (thank you Raymond for inviting me to visit you in Lingnan University). The subjects of my communications, in Shenzhen and in Hong Kong, were not so important, as Eugene and Raymond wrote to me.

I guess that Gauss-Seidel and power iterative methods are utilized in solving Data Science problems. In this talk I'll share with you two simple remarks, related to such methods.

First remark: I had to show to my students a system $A\mathbf{x} = \mathbf{f}$ solvable by Gauss-Seidel (G.S.) method even if A does not satisfy the standard sufficient conditions for G.S. convergence ...

Definition of Gauss-Seidel iteration matrix:

$$\begin{array}{l} A \in \mathbb{C}^{n\times n} \\ a_{i,i} \neq 0\,\forall i \end{array} \qquad \to \qquad A_{GS} = -\begin{bmatrix} a_{1,1} & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & 0 \\ a_{n,1} & \cdot & \cdot & a_{n,n} \end{bmatrix}^{-1} \begin{bmatrix} 0 & a_{1,2} & \cdot & a_{1,n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & a_{n-1,n} \\ 0 & \cdot & \cdot & 0 \end{bmatrix}$$

[Recall: Given any $\mathbf{f} \in \mathbb{C}^n$, and set $\mathbf{v_f} = \begin{bmatrix} a_{1,1} & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & 0 \\ a_{n,1} & \cdot & \cdot & a_{n,n} \end{bmatrix}^{-1} \mathbf{f}$, we have that

$$A^{-1}\mathbf{f} \text{ solves the equation } \mathbf{x} = A_{GS}\mathbf{x} + \mathbf{v_f}$$

and the sequence $\mathbf{x}_{k+1} = A_{GS}\mathbf{x}_k + \mathbf{v_f}$, $k = 0, 1, \ldots$, converges to $A^{-1}\mathbf{f}$, $\forall\, \mathbf{x}_0 \in \mathbb{C}^n$, if and only if $\rho(A_{GS}) < 1$]

We know that: $$\begin{bmatrix} y & \mathbf{0}^T \\ \mathbf{w} & Z \end{bmatrix}^{-1} = \begin{bmatrix} 1/y & \mathbf{0}^T \\ -\frac{1}{y}Z^{-1}\mathbf{w} & Z^{-1} \end{bmatrix}, \quad y \neq 0,\ \det Z \neq 0\,.$$

$$\Rightarrow \quad A_{GS} = -\begin{bmatrix} 1/a_{11} & \mathbf{0}^T \\ -\frac{1}{a_{11}}\begin{bmatrix} a_{22} & 0 \cdot & 0 \\ \cdot & \cdot & 0 \\ a_{n2} & \cdot & a_{n,n} \end{bmatrix}^{-1}\begin{bmatrix} a_{21} \\ \cdot \\ a_{n1} \end{bmatrix} & \begin{bmatrix} a_{22} & 0 \cdot & 0 \\ \cdot & \cdot & 0 \\ a_{n2} & \cdot & a_{n,n} \end{bmatrix}^{-1} \end{bmatrix} \begin{bmatrix} 0 & a_{12} \quad \cdots \quad \cdot \quad \cdots \quad a_{1n} \\ \mathbf{0} & \begin{bmatrix} 0 & a_{23} & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & a_{n-1,n} \\ 0 & \cdot & 0 & 0 \end{bmatrix} \end{bmatrix}$$

$$= -\begin{bmatrix} 0 & a_{12}/a_{11} \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad a_{1n}/a_{11} \\ \mathbf{0} & -\frac{1}{a_{11}}\begin{bmatrix} a_{22} & 0 \cdot & 0 \\ \cdot & \cdot & 0 \\ a_n 2 & \cdot & a_{nn} \end{bmatrix}^{-1}\begin{bmatrix} a_{21} \\ \cdot \\ a_{n1} \end{bmatrix}(a_{12} \quad \cdots \quad a_{1n}) + B_{GS} \end{bmatrix}$$

Let B be the lower right $(n-1) \times (n-1)$ submatrix of A:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & a_{1,n} \\ a_{2,1} & & & \\ \cdot & & B & \\ a_{n,1} & & & \end{bmatrix} \quad \text{and consider} \quad B_{GS} = - \begin{bmatrix} a_{2,2} & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & 0 \\ a_{n,2} & \cdot & \cdot & a_{n,n} \end{bmatrix}^{-1} \begin{bmatrix} 0 & a_{2,3} & \cdot & a_{2,n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & a_{n-1,n} \\ 0 & \cdot & \cdot & 0 \end{bmatrix}$$

THEOREM: $$\begin{bmatrix} a_{21} \\ a_{31} \\ \cdot \\ a_{n1} \end{bmatrix} = \alpha \begin{bmatrix} a_{22} \\ a_{32} \\ \cdot \\ a_{n2} \end{bmatrix}, \ \alpha = \frac{a_{21}}{a_{22}} \quad \Rightarrow \quad \rho(A_{GS}) = \max\{\left|\frac{a_{12}a_{21}}{a_{11}a_{22}}\right|, \rho(B_{GS})\},$$

i.e. $\rho(A_{GS})$ can be written in terms of $\rho(B_{GS})$.

$\rightarrow$ Examples of linear systems $A\mathbf{x} = \mathbf{f}$ solvable by Gauss-Seidel method, where A does not satisfy the standard sufficient conditions for its convergence (A is not positive definite, not with strictly dominant diagonal, not irreducible with weakly dominant diagonal):

$$A = \begin{bmatrix} a_{1,1} & a_{12} & \cdot & a_{1,n} \\ \alpha a_{22} & & & \\ \cdot & & B & \\ \alpha a_{n2} & & & \end{bmatrix}, \quad \alpha = \frac{a_{21}}{a_{22}}, \; \rho(B_{GS}) < 1, \; \left| \frac{a_{12}a_{21}}{a_{11}a_{22}} \right| < 1 \Rightarrow \; \rho(A_{GS}) < 1$$

(independently from the values of $a_{1,i}$, $i = 3, \ldots, n$). For instance,

$$A = \begin{bmatrix} -1 & 3.8 & \frac{1}{2} & -7 \\ \frac{1}{2} & & & \\ -\frac{1}{4} & & B & \\ 0 & & & \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 & 0 \\ -1 & \pm 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \Rightarrow \alpha = \frac{1}{4}, \; \rho(B_{GS}) < 1, \; \left| \frac{3.8 * \frac{1}{2}}{-1 * 2} \right| < 1$$

$$\Rightarrow \quad \rho(A_{GS}) < 1$$

$\rightarrow$ Assume, moreover, that

$$B = \begin{bmatrix} a_{2,2} & a_{2,3} & \cdot & a_{2,n} \\ a_{3,2} & & & \\ \cdot & & C & \\ a_{n,2} & & & \end{bmatrix} \text{ satisfies } \begin{bmatrix} a_{32} \\ a_{42} \\ \cdot \\ a_{n2} \end{bmatrix} = \beta \begin{bmatrix} a_{33} \\ a_{43} \\ \cdot \\ a_{n3} \end{bmatrix}, \ \beta = \frac{a_{32}}{a_{33}},$$

then, by the THEOREM,

$$\rho(A_{GS}) = \max\{\left|\frac{a_{12}a_{21}}{a_{11}a_{22}}\right|, \rho(B_{GS})\} = \max\left\{\left|\frac{a_{12}a_{21}}{a_{11}a_{22}}\right|, \max\{\left|\frac{a_{23}a_{32}}{a_{22}a_{33}}\right|, \rho(C_{GS})\} \right\}$$

$$= \max\{\left|\frac{a_{12}a_{21}}{a_{11}a_{22}}\right|, \left|\frac{a_{23}a_{32}}{a_{22}a_{33}}\right|, \rho(C_{GS})\}$$

$\rightarrow$ $\cdots$

$\rightarrow$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ \alpha a_{22} & a_{22} & a_{23} & a_{24} \\ \alpha\beta a_{33} & \beta a_{33} & a_{33} & a_{34} \\ \alpha\beta\gamma a_{44} & \beta\gamma a_{44} & \gamma a_{44} & a_{44} \end{bmatrix}, \ \alpha = \frac{a_{21}}{a_{22}}, \ \beta = \frac{a_{32}}{a_{33}}, \ \gamma = \frac{a_{43}}{a_{44}}$$

$$\Rightarrow \quad \rho(A_{GS}) = \max\{\left|\frac{a_{12}a_{21}}{a_{11}a_{22}}\right|, \left|\frac{a_{23}a_{32}}{a_{22}a_{33}}\right|, \left|\frac{a_{34}a_{43}}{a_{33}a_{44}}\right|\}$$

(in fact, $C_{GS} = \begin{bmatrix} 0 & -a_{34}/a_{33} \\ 0 & a_{34}a_{43}/a_{33}a_{44} \end{bmatrix}$),

A_{GS} is upper triangular with $[A_{GS}]_{ii} = \dfrac{a_{i-1,i}a_{i,i-1}}{a_{i-1,i-1}a_{i,i}}$, $i = 2, 3, 4$.

The case n generic is analogous.

$\rightarrow$

$$A = \begin{bmatrix} a_0 & a_{-1} & a_{-2} & a_{-3} \\ \alpha a_0 & a_0 & a_{-1} & a_{-2} \\ \alpha^2 a_0 & \alpha a_0 & a_0 & a_{-1} \\ \alpha^3 a_0 & \alpha^2 a_0 & \alpha a_0 & a_0 \end{bmatrix}, \ \alpha = \frac{a_1}{a_0} \quad \Rightarrow \quad \rho(A_{GS}) = \left| \frac{a_1 a_{-1}}{(a_0)^2} \right|$$

A_{GS} is upper triangular Toeplitz with $[A_{GS}]_{ii} = \dfrac{a_1 a_{-1}}{(a_0)^2}$, $i = 2, 3, 4$.

For example, if $a_{-1} = 0$ then Gauss-Seidel converges in 4 steps

The case n generic is analogous.

The THEOREM is a corollary of the following more general result:

Let B and C be respectively the lower right $(n-r)\times(n-r)$ and the upper left $r\times r$ submatrices of A,

$$A = \begin{bmatrix} C & R \\ \\ S & B \end{bmatrix} \in \mathbb{C}^{n\times n}, \quad C \in \mathbb{C}^{r\times r}, \quad \text{and consider} \quad B_{GS} \quad \text{and} \quad C_{GS}. \ :$$

$$\begin{array}{c} S\mathbf{e}_i = \alpha_i B\mathbf{e}_1, \ \alpha_i = \dfrac{a_{r+1,i}}{a_{r+1,r+1}}, \ i=1,\dots,r, \\ R\mathbf{e}_1 = \gamma C\mathbf{e}_1, \ \gamma = \dfrac{a_{1,r+1}}{a_{11}} \end{array} \quad \Rightarrow \quad \rho(A_{GS}) = \max\{\rho(C_{GS}), \left|\frac{a_{1,r+1}a_{r+1,1}}{a_{11}a_{r+1,r+1}}\right|, \rho(B_{GS})\}$$

In other words, if the columns of S are all proportional to the first column of B and the first column of R is proportional to the first column of C, then $\rho(A_{GS})$ can be written in terms of $\rho(B_{GS})$ and $\rho(C_{GS})$.

Proof: use the equality

$$\begin{bmatrix} Y & O \\ W & Z \end{bmatrix}^{-1} = \begin{bmatrix} Y^{-1} & O \\ -Z^{-1}WY^{-1} & Z^{-1} \end{bmatrix}, \quad \det Y \neq 0,\ \det Z \neq 0$$

Example:

$$A = \left[\begin{array}{ccc|cccc} & & & 1/4 & \# & \# & \# \\ & C & & 1/8 & \# & \# & \# \\ & & & 1/12 & \# & \# & \# \\ -4 & 1 & 6 & & & & \\ 2 & -\frac{1}{2} & -3 & & B & & \\ 0 & 0 & 0 & & & & \\ 0 & 0 & 0 & & & & \end{array}\right], \quad B = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix},$$

$$\alpha_1 = -2,\ \alpha_2 = \tfrac{1}{2},\ \alpha_3 = 3, \qquad \gamma = \tfrac{1}{4}$$

$$\rho(B_{GS}) < 1, \quad \rho(C_{GS}) < 1, \quad \left|\frac{-4 * 1/4}{1 * 2}\right| < 1 \quad \Rightarrow \quad \rho(A_{GS}) < 1$$

(independently from the values of the entries #)

Second remark:

Problem: Given $\mathcal{X} \in \mathbb{C}^{n\times n}$ invertible which transforms $A \in \mathbb{C}^{n\times n}$ into its Jordan canonical form J_A ($\mathcal{X}^{-1}A\mathcal{X} = J_A$), find $\mathcal{Y} \in \mathbb{C}^{n\times n}$ invertible which transforms A^{-1} into its Jordan canonical form $J_{A^{-1}}$ ($\mathcal{Y}^{-1}A^{-1}\mathcal{Y} = J_{A^{-1}}$).

[Motivation: Let $A \in \mathbb{C}^{n\times n}$ with spectrum $\sigma(A) = \{\mu_1, \dots, \mu_n\} : 0 < |\mu_n| < |\mu_j|$, $\forall\, \mu_j \neq \mu_n$, $m_a(\mu_n) = m_g(\mu_n)$. The study of the convergence of the vectors $\mu_n^k A^{-k}\mathbf{v}$ generated by the power method applied to A^{-1} is simplified, in the case A not diagonalizable, if the initial vector $\mathbf{v}$ is represented as a linear combination of the columns of $\mathcal{Y}$, rather than of the columns of $\mathcal{X}$.

Moreover, the study of the relation between $\mathcal{Y}$ and $\mathcal{X}$ makes more possible a comparison between the rate of convergence of the sequence $\mu_n^k A^{-k}\mathbf{v}$ and the rate of convergence of the sequence $(1/\mu_1)^k A^k\mathbf{v}$, in case the further conditions $0 < |\mu_n| < |\mu_j| < |\mu_1|$, $\forall\, \mu_j \neq \mu_n, \mu_1$, $m_a(\mu_1) = m_g(\mu_1)$ are satisfied.]

Solution of the Problem: $\mathcal{Y} = \mathcal{X}\mathcal{M}$ where $\mathcal{M}$ is a block diagonal matrix (with the same pattern of J_A and $J_{A^{-1}}$) whose diagonal blocks M are *upper triangular Tartaglia-Toeplitz matrices* (see the following slides)

Assume that for $A \in \mathbb{C}^{n\times n}$ invertible, there exist $X \in \mathbb{C}^{n\times s}$ $(s \leq n)$, with columns linearly independent, and $\mu \in \mathbb{C}$, $\mu \neq 0$, such that

$$AX = X(\mu I_s + Z_s^T) = X \left.\begin{bmatrix} \mu & 1 & 0 & . \\ 0 & \mu & . & 0 \\ . & . & . & 1 \\ 0 & . & 0 & \mu \end{bmatrix}\right\} s \qquad \leftarrow \text{ definition of } Z_s^T$$

(such assumption is verified by any Jordan block of A). Then $A^{-1}X = X(\mu I_s + Z_s^T)^{-1}$, and therefore the following result holds:

$$M \in \mathbb{C}^{s\times s} \ : \ (\mu I_s + Z_s^T)^{-1} M = M(\frac{1}{\mu} I_s + Z_s^T) \qquad (\#)$$

$$\Rightarrow \qquad A^{-1}Y = Y(\frac{1}{\mu} I_s + Z_s^T) = Y \begin{bmatrix} \frac{1}{\mu} & 1 & 0 & . \\ 0 & \frac{1}{\mu} & . & 0 \\ . & . & . & 1 \\ 0 & . & 0 & \frac{1}{\mu} \end{bmatrix}, \text{ with } Y = XM.$$

So, the Problem is solved if we find a matrix M satisfying $(\#)$ and invertible.

Conditions equivalent to condition (#):

$$M \in \mathbb{C}^{s\times s} \ : \ (\mu I_s + Z_s^T)^{-1} M = M(\frac{1}{\mu} I_s + Z_s^T) \qquad (\#)$$

if and only if

$$M = (\mu I_s + Z_s^T) M (\frac{1}{\mu} I_s + Z_s^T)$$

if and only if

$$M = (I_s + \frac{1}{\mu} Z_s^T) M (I_s + \mu Z_s^T)$$

if and only if

$$O = \mu M Z_s^T + \frac{1}{\mu} Z_s^T M + Z_s^T M Z_s^T$$

One invertible matrix satisfying (#):

The upper triangular Tartaglia matrix P and the lower shift matrix Z satisfy the identity

$$(I+Z^T)P(I-Z^T)=P, \qquad P=\begin{bmatrix} 1 & 1 & 1 & 1 & \cdot \\ 0 & 1 & 2 & 3 & \cdot \\ 0 & 0 & 1 & 3 & \cdot \\ 0 & 0 & 0 & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}, \quad Z=\begin{bmatrix} 0 & 0 & 0 & 0 & \cdot \\ 1 & 0 & 0 & 0 & \cdot \\ 0 & 1 & 0 & 0 & \cdot \\ 0 & 0 & 1 & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Moreover, for $\mu\in\mathbb{C}$, $\mu\neq 0$, and $d((\pm\mu)^j) = \operatorname{diag}((\pm\mu)^j,\ j=0,1,\ldots)$, we have

$$\begin{aligned} d(\mu^j)(I+Z^T) &= (I+\tfrac{1}{\mu}Z^T)d(\mu^j), \\ (I-Z^T)d((-\mu)^j) &= d((-\mu)^j)(I+\mu Z^T), \end{aligned}$$

$$\Rightarrow\ d(\mu^j)Pd((-\mu)^j) = d(\mu^j)(I+Z^T)P(I-Z^T)d((-\mu)^j) = (I+\frac{1}{\mu}Z^T)d(\mu^j)Pd((-\mu)^j)(I+\mu Z^T),$$

i.e. (1) $\underline{\Gamma=\Gamma_s := d_s(\mu^j)P_s d_s((-\mu)^j)}$ is an invertible matrix that satisfies (#)

Note: Γ is an upper triangular matrix with first row $[1\,|-\mu\,|\,\mu^2\,|-\mu^3\,|\,\cdots\,|\,(-\mu)^{s-1}]$.

Look for other matrices satisfying (#):

There could be invertible matrices M satisfying (#) *better* than Γ, for example such that the computation of the matrix-vector product XM requires less arithmetic operations than the computation of $X\Gamma$. This question leads to the study of the space

$$\mathcal{L} = \mathcal{L}_s = \{M \in \mathbb{C}^{s\times s} : \ (\#) \text{ holds}\}$$

<u>Remark</u>: If $M \in \mathcal{L}_{s-1}$ and $\mathbf{e}_1^T M = [1 \,|\, -\mu \,|\, \mu^2 \,|\, -\mu^3 \,|\, \cdots \,|\, (-\mu)^{s-2}]$, then $\begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & -\mu^2 M \end{bmatrix} \in \mathcal{L}_s$.

Γ_{s-1} satisfies the hypotheses in the Remark, thus we have the following other example of invertible matrix in $\mathcal{L}$:

$$(2) \quad \underline{G = G_s := \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & -\mu^2\Gamma_{s-1} \end{bmatrix}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & . & 0 \\ 0 & -\mu^2 & \mu^3 & -\mu^4 & \mu^5 & . & (-1)^{s-1}\mu^s \\ 0 & 0 & \mu^4 & -2\mu^5 & 3\mu^6 & . & (-1)^{s-1}(s-2)\mu^{s+1} \\ 0 & . & 0 & -\mu^6 & 3\mu^7 & . & . \\ 0 & . & . & 0 & \mu^8 & . & . \\ . & . & . & . & . & . & (-1)^{s-1}(s-2)\mu^{2s-3} \\ 0 & 0 & 0 & 0 & 0 & . & (-1)^{s-1}\mu^{2s-2} \end{bmatrix} \in \mathcal{L}$$

Upper triangular Toeplitz matrices allow to describe $\mathcal{L}$:

Condition (#) holds if and only if $M = (\mu I + Z^T)M(\frac{1}{\mu}I + Z^T)$ iff

$$\mu M Z^T + \frac{1}{\mu} Z^T M + Z^T M Z^T = O\,. \qquad (\#')$$

Observe that if M satisfies $(\#')$ and T is any upper triangular Toeplitz matrix, then also the matrix MT satisfies $(\#')$ (upper triangular Toeplitz matrices commute with Z^T!). Thus

(3) $\underline{\{GT :\ T = \text{upper triangular Toeplitz}\} \subset \mathcal{L}}$.

Viceversa, if $M = (I + \frac{1}{\mu}Z^T)M(I + \mu Z^T)$, then, for any invertible $L \in \mathcal{L}$, we have

$$L^{-1}M = L^{-1}(I + \frac{1}{\mu}Z^T)M(I + \mu Z^T) = (I + \mu Z^T)^{-1}L^{-1}M(I + \mu Z^T)$$

i.e. $(I+\mu Z^T)L^{-1}M = L^{-1}M(I+\mu Z^T)$. But the latter identity implies that $L^{-1}M$ must be upper triangular Toeplitz, that is, M must be in the space $\{LT :\ T = \text{upper triangular Toeplitz}\}$.

$\rightarrow$ **Characterization of $\mathcal{L} = \{M \in \mathbb{C}^{s\times s} :\ (\#) \text{ holds}\}$:**

---

(4) $\mathcal{L}$ is a space of upper triangular Tartaglia-Toeplitz matrices, precisely

$$\mathcal{L} = \mathcal{L}_s = \{GT :\ T = \text{upper triangular Toeplitz}\}$$
$$= \operatorname{Span}\{G,\, GZ^T,\, G(Z^T)^2, \dots,\, G(Z^T)^{s-1}\}.$$

Note that $\mathbf{e}_1^T G(Z^T)^{k-1} = \mathbf{e}_k^T$, $k = 1, \dots, s$, and therefore the generic matrix of $\mathcal{L}$ is determined by its first row.

As a consequence of (4), the matrix Γ itself must be of type GT, for some upper triangular Toeplitz T. By imposing $[1 \,|\, -\mu \,|\, \mu^2 \,|\, -\mu^3 \,|\, \cdots \,|\, (-\mu)^{s-1}] = \mathbf{e}_1^T\Gamma = \mathbf{e}_1^T GT = \mathbf{e}_1^T T$, we guess that $T = (I + \mu Z^T)^{-1}$:

(5) $\underline{\Gamma = G(I + \mu Z^T)^{-1}}$

Recall that Γ is defined in terms of P_s, the upper triangular Tartaglia matrix of order s, whereas G is defined in terms of P_{s-1}, the upper triangular Tartaglia matrix of order $s-1$. Thus, (5) yields an equality involving such matrices, as well as an equality involving their inverses:

$$
(6)\quad P_k = \begin{bmatrix} 1 & \begin{matrix} 1\;1\;1 \cdots\cdots\cdots 1\;1 \end{matrix} \\ \mathbf{0} & P_{k-1} \begin{bmatrix} 1 & 1 & 1 & . & 1 \\ 0 & 1 & 1 & . & 1 \\ . & . & 1 & . & . \\ 0 & . & . & . & 1 \\ 0 & . & . & 0 & 1 \end{bmatrix} \end{bmatrix} \quad \Rightarrow \quad P_k^{-1} = \begin{bmatrix} 1 & -[1\;0\;0 \cdots\cdots\cdots 0\;0]P_{k-1}^{-1} \\ \mathbf{0} & \begin{bmatrix} 1 & -1 & 0 & . & 0 \\ 0 & 1 & -1 & . & . \\ . & . & . & . & 0 \\ . & . & . & . & -1 \\ 0 & . & . & 0 & 1 \end{bmatrix} P_{k-1}^{-1} \end{bmatrix}
$$

where P_k and P_k^{-1} are the upper-left $k \times k$ submatrices of P and P^{-1}, with

$$
P_{ij} = \binom{j-1}{i-1}, \quad [P^{-1}]_{ij} = (-1)^{j-i}\binom{j-1}{i-1} \quad \text{if } i \leq j\,.
$$

Observe that the equalities in (6) can be used to compute matrix-vector products where the matrix is P_s, or to solve linear systems with P_s as coefficient matrix.

For example, the repeated application of the representation of P_k^{-1}, for $k = s, s-1, \ldots, 2$, yields a procedure for computing $P_s^{-1}\mathbf{f}$, $\mathbf{f} \in \mathbb{C}^s$, which avoids the computation of the entries of P_s^{-1} and requires only $s(s-1)/2$ additive operations.

Here below is the detailed procedure:

$$
\begin{array}{l}
\text{input: } f_1, f_2, \ldots, f_s \\
i = 1 \\
\bullet_1 = f_s \\
\text{for } i = 1, \ldots, s-1 \text{ do } \{ \\
\quad z = \bullet_1 \ ; \\
\quad app(1) = \bullet_1 \ ; \\
\quad \bullet_1 = f_{s-i} - z \ ; \ \bullet_{i+1} = 0 \ ; \\
\quad \text{for } j = 2, \ldots, i+1 \text{ do } \{ \\
\qquad app(j) = \bullet_j \ ; \\
\qquad \bullet_j = app(j-1) - app(j); \\
\quad \} \\
\}
\end{array}
$$

Open problem: what is the matrix M in $\mathcal{L}_s = \{GT : \ T = \text{upper triangular Toeplitz}\}$ for which the computation of $Y = XM$, $X \in \mathbb{C}^{n \times s}$, requires the minimum number of arithmetic operations ?