Lecture V Partitioned Matrices

Lei Du
dulei@dlut.edu.cn

School of Mathematical Sciences
Dalian University of Technology

Octomber, 2022

Contents

1. Introduction
2. Submatrices

   blocked

3. Matrix products and partitioned matrices
4. Special cases of partitioned matrices
5. Determinants of partitioned matrices

1 Introduction

2 Submatrices

3 Matrix products and partitioned matrices

4 Special cases of partitioned matrices

5 Determinants of partitioned matrices

Introduction

Partitioned matrices, also called block matrices, play an important role in matrix and tensor calculus. They are usually employed for the computation of matrix products, especially for Khatri-Rao and Kronecker products. Partitioned matrices are also underlying the definition of certain structured matrices such as Hamiltonian, Hadamard, Fourier, Toeplitz, and Hankel matrices, and subsequently block-Toeplitz and block-Hankel matrices.

This chapter has several objectives:

- to define the notions of submatrices and partitioned matrices
- to describe examples of partitioned matrices for the computation of matrix products
- to present a few special cases such as block-diagonal matrices, Jordan forms, block-triangular matrices, block-Toeplitz and Hankel matrices
- to define block operations such as transposition, trace, addition and multiplication, as well as the determinants, and the ranks of certain partitioned matrices
- to introduce elementary operations and associated matrices, used for block triangularization, block-diagonalization, block-factorization, block- inversion, and generalized inversion of 2×2 block matrices
- to use inversion formulae of block matrices to deduce several fundamental results such as the matrix inversion lemma, the inversion of a partitioned Gram matrix, and recursive inversion with respect to the order of a square partitioned matrix
- matrix representations of a linear map and a bilinear/sesquilinear form, Quadratic forms and Hermitian forms
- eigenvalue and eigenvector, and generalized eigenvalue problem
- to provide an example of application of the recursive inversion formula of a 2×2 block matrix, for demonstrating the Levinson algorithm which is an algorithm widely used in signal processing for the estimation of the parameters of an autoregressive (AR) model,

1 Introduction

2 Submatrices

3 Matrix products and partitioned matrices

4 Special cases of partitioned matrices

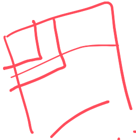
5 Determinants of partitioned matrices

Submatrices

A submatrix $\mathbf{B}\left(m_i, n_j\right)$ of a matrix $\mathbf{A}(m, n)$, with $m_i \leq m$ and $n_j \leq n$, is a matrix whose elements are positioned at the intersections of the m_i rows and n_j columns of A defined by the sets of indices:

$$\alpha_{m_i} = \{i_k, k \in \langle m_i \rangle\} \subseteq \langle m \rangle, \beta_{n_j} = \{j_l, l \in \langle n_j \rangle\} \subseteq \langle n \rangle.$$

Thus, the element $a_{i_k j_l}$ of $\mathbf{A}$ is given by $a_{i_k j_l} = \left(\mathbf{e}_{i_k}^{(m)}\right)^T \mathbf{A}\mathbf{e}_{j_l}^{(n)}$. Subsequently, by defining the row and column selection matrices:



$$\mathbf{M} = \left[\mathbf{e}_{i_1}^{(m)}, \cdots, \mathbf{e}_{i_{m_i}}^{(m)}\right] \text{ and } \mathbf{N} = \left[\mathbf{e}_{j_1}^{(n)}, \cdots, \mathbf{e}_{j_{n_j}}^{(n)}\right],$$

we can write $\mathbf{B}\left(m_i, n_j\right)$ as:

A(1:3, 2:4)

$$\mathbf{B}\left(m_i, n_j\right) = \mathbf{M}^T\mathbf{A}\mathbf{N} = \begin{bmatrix} a_{i_1, j_1} & \cdots & a_{i_1, j_{n_j}} \\ \vdots & & \vdots \\ a_{i_{m_i}, j_1} & \cdots & a_{i_{m_i}, j_{n_j}} \end{bmatrix}.$$

In the case of a square matrix $\mathbf{A}$ of order n, a principal submatrix of order r is a submatrix $\mathbf{B}(r, r)$ whose elements are positioned at the intersections of the same set of r rows and r columns, that is, defined by the same set of indices $\alpha_r = \{i_k, k \in \langle r \rangle\} \subseteq \langle n \rangle$. A principal submatrix of order r contains r elements of the main diagonal of $\mathbf{A}$. There are $C_n = \frac{n!}{r!(n-r)!}$ principal submatrices of order r.

Partitioned matrices

Let $\{\alpha_{m_1}, \cdots, \alpha_{m_R}\}$ and $\{\beta_{n_1}, \cdots, \beta_{ns}\}$ be partitions of the sets $\{1, \cdots, m\}$ and $\{1, \cdots, n\}$, respectively, with $m_r \in \langle m \rangle$ and $n_s \in \langle n \rangle$, such that $\sum_{r=1}^{R} m_r = m$ and $\sum_{s=1}^{S} n_s = n$. It is said that matrices $\mathbf{A}_{rs}$ of dimensions $(m_{r+} n_s)$ form a partition of the matrix $\mathbf{A} \in \mathbb{K}^{m \times n}$ into (R, S) blocks, or that $\mathbf{A}$ is partitioned into (R, S) blocks, if $\mathbf{A}$ can be written as:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1S} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{R1} & \mathbf{A}_{R2} & \cdots & \mathbf{A}_{RS} \end{bmatrix} = [\mathbf{A}_{rs}], r \in \langle R \rangle, s \in \langle S \rangle. \tag{1}$$

Such a partitioning with blocks of different dimensions is said to be unbalanced.

Partitioned matrices

The submatrix $\mathbf{A}_{rs}$ can be expressed as:

$$\mathbf{A}_{rs} = \begin{bmatrix} a_{m_1+\cdots+m_{r-1}+1, n_1+\cdots+n_{x-1}+1} & \cdots & a_{m_1+\cdots+m_{r-1}+1, n_1+\cdots+n_{x-1}+n_x} \\ \vdots & \ddots & \vdots \\ a_{m_1+\cdots+m_{r-1}+m_r, n_1+\cdots+n_{s-1}+1} & \cdots & a_{m_1+\cdots+m_{n-1}+m_{r,n_1+\cdots+n_{s-1}+n_s}} \end{bmatrix}$$
$$\in \mathbb{K}^{m_r \times n_s}.$$

All submatrices of the same row-block (r) contain the same number (m_r) of rows. Similarly, all submatrices of the same column-block (s) contain the same number (n_s) of columns, that is:

$$\begin{bmatrix} \mathbf{A}_{r1} & \mathbf{A}_{r2} & \cdots & \mathbf{A}_{rS} \end{bmatrix} \in \mathbb{K}^{m_r \times n}, \begin{bmatrix} \mathbf{A}_{1s} \\ \mathbf{A}_{2s} \\ \vdots \\ \mathbf{A}_{Rs} \end{bmatrix} \in \mathbb{K}^{m \times n_s}.$$

It is then said that the submatrices $\mathbf{A}_{rs}$ are of compatible dimensions.

Partitioned matrices

In the particular case where $n = 1$, the partitioned matrix (1) becomes a blockcolumn vector:

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_R \end{bmatrix} \in \mathbb{K}^{m\times 1}, \mathbf{a}_r \in \mathbb{K}^{m_r\times 1}, r \in \langle R\rangle.$$

Similarly, when $m = 1$, the partitioned matrix (1) becomes a block-row vector:

$$\mathbf{a}^T = \left[\mathbf{a}_1^T\, \mathbf{a}_2^T\, \cdots\, \mathbf{a}_S^T\right] \in \mathbf{K}^{1\times n}, \mathbf{a}_s \in \mathbb{K}^{n_s\times 1}, s \in \langle S\rangle.$$

If all the blocks $\mathbf{A}_{rs}$ have the same dimensions $P \times Q$, that is, when $m_r = P, \forall r \in \langle R\rangle$, and $n_s = Q, \forall s \in \langle S\rangle$, then the space of partitioned matrices into (R, S) blocks, with entries in the space $\mathbb{K}^{P\times Q}$ (also written $\mathcal{M}_{P\times Q}(\mathbb{K})$), will be denoted $\mathcal{M}_{R\times S}(\mathcal{M}_{P\times Q}(\mathbb{K}))$. The partitioning is then said to be balanced.

1 Introduction

2 Submatrices

3 **Matrix products and partitioned matrices**

4 Special cases of partitioned matrices

5 Determinants of partitioned matrices

Matrix products

Given two rectangular matrices $\mathbf{A} \in \mathbb{K}^{I \times J}$ and $\mathbf{B} \in \mathbb{K}^{J \times K}$, the product $\mathbf{C} = \mathbf{AB} \in \mathbb{K}^{I \times K}$ can be written in terms of matrices partitioned into column blocks or row blocks:

$$\mathbf{AB} = [\mathbf{AB}_{.1}, \mathbf{AB}_{.2}, \cdots, \mathbf{AB}_{.K}] = \begin{bmatrix} \mathbf{A}_{1}\mathbf{B} \\ \mathbf{A}_{2.}\mathbf{B} \\ \vdots \\ \mathbf{A}_{I}\mathbf{B} \end{bmatrix}.$$

Two matrix products play an important role in matrix calculation. These are the Kronecker and Khatri-Rao products.

Vector Kronecker product

Let $\mathbf{u} \in \mathbb{K}^I$ and $\mathbf{v} \in \mathbb{K}^J$. Their Kronecker product is defined as:

$$\mathbf{x} = \mathbf{u} \otimes \mathbf{v} = \begin{bmatrix} u_1 \mathbf{v} \\ \vdots \\ u_I \mathbf{v} \end{bmatrix} \in \mathbb{K}^{IJ}$$
$$= \left[u_1 v_1, u_1 v_2, \cdots, u_1 v_J, u_2 v_1, \cdots, u_I v_J\right]^T.$$

This is a vector partitioned into I blocks of dimension J. The element $u_i v_j$ is positioned at position $j + (i-1)J$.

Matrix Kronecker product

Given $\mathbf{A} \in \mathbb{K}^{I \times J}$ and $\mathbf{B} \in \mathbb{K}^{M \times N}$, the Kronecker product to the right of $\mathbf{A}$ by $\mathbf{B}$ is the matrix $\mathbf{C} \in \mathbb{K}^{IM \times JN}$ defined as :

$$\mathbf{C} = \mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1J}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2J}\mathbf{B} \\ \vdots & \vdots & & \vdots \\ a_{I1}\mathbf{B} & a_{I2}\mathbf{B} & \cdots & a_{IJ}\mathbf{B} \end{bmatrix} = [a_{ij}\mathbf{B}] \, .$$

This is a matrix partitioned into (I, J) blocks, the block (i, j) being the matrix $a_{ij}\mathbf{B} \in \mathbb{K}^{M \times N}$. The element $a_{ij}b_{mn}$ is positioned at position $((i-1)M + m, (j-1)N + n)$ in $\mathbf{A} \otimes \mathbf{B}$.

Matrix Kronecker product

The j th column-block of $\mathbf{A} \otimes \mathbf{B}$ is given by:

$$\mathbf{A}_{.j} \otimes \mathbf{B} = \begin{bmatrix} a_{1j}\mathbf{B} \\ \vdots \\ a_{Ij}\mathbf{B} \end{bmatrix} = \left[\mathbf{A}_{.j} \otimes \mathbf{B}_{.1} \mathbf{A}_{.j} \otimes \mathbf{B}_{.2} \cdots \mathbf{A}_{.j} \otimes \mathbf{B}_{.N}\right], j \in \langle J \rangle$$

Subsequently, the columns of $\mathbf{A} \otimes \mathbf{B}$ are composed of all the Kronecker products of a column of $\mathbf{A}$ with a column of $\mathbf{B}$, the columns being taken in lexicographical order. Similarly, $\mathbf{A} \otimes \mathbf{B}$ can be decomposed into I row-blocks $\mathbf{A}_{i.} \otimes \mathbf{B}$, with $i \in \langle I \rangle$, the IM rows being composed of all the Kronecker products of a row of $\mathbf{A}$ with a row of $\mathbf{B}$. Therefore, $\mathbf{A} \otimes \mathbf{B}$ can be broken into blocks such that:

$$\begin{aligned} \mathbf{A} \otimes \mathbf{B} &= [\mathbf{A}_{.1} \otimes \mathbf{B} \cdots \mathbf{A}_{.J} \otimes \mathbf{B}] \\ &= [\mathbf{A}_{.1} \otimes \mathbf{B}_{.1} \cdots \mathbf{A}_{.1} \otimes \mathbf{B}_{.N} \cdots \mathbf{A}_{.J} \otimes \mathbf{B}_{.1} \cdots \mathbf{A}_{.J} \otimes \mathbf{B}_{.N}] \\ &= \begin{bmatrix} \mathbf{A}_{1.} \otimes \mathbf{B} \\ \vdots \\ \mathbf{A}_{I.} \otimes \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{1.} \otimes \mathbf{B}_{1.} \\ \vdots \\ \mathbf{A}_{1.} \otimes \mathbf{B}_{M.} \\ \vdots \\ \mathbf{A}_{I.} \otimes \mathbf{B}_{1.} \\ \vdots \\ \mathbf{A}_{I.} \otimes \mathbf{B}_{M.} \end{bmatrix}. \end{aligned}$$

Matrix Kronecker product

The Kronecker product can be used to write the matrix $\mathbf{A}$ partitioned into (R, S) blocks, defined in (1), as follows:

$$\mathbf{A} = \sum_{r=1}^{R} \sum_{s=1}^{S} \mathbf{E}_{rs}^{(R \times S)} \otimes \mathbf{A}_{rs}$$

where $\mathbf{E}_{rs}^{(R \times S)}$, for $r \in \langle R \rangle$ and $s \in \langle S \rangle$, are the matrices of the canonical basis of the space $\mathbb{K}^{R \times S}$, that is, with the (r, s) th element equal to 1 and all others equal to zero.

例

For $R = 2$ and $S = 3$, we have:

$$\mathbf{A} = \sum_{r=1}^{2} \sum_{s=1}^{3} \mathbf{E}_{rs}^{(2 \times 3)} \otimes \mathbf{A}_{rs} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{bmatrix}.$$

Khatri-Rao product

Given $\mathbf{A} \in \mathbb{K}^{I \times J}$ and $\mathbf{B} \in \mathbb{K}^{K \times J}$ having the same number of columns, the Khatri-Rao product of $\mathbf{A}$ with $\mathbf{B}$, denoted by $\mathbf{A} \diamond \mathbf{B} \in \mathbb{K}^{IK \times J}$, is defined as:

$$\mathbf{A} \diamond \mathbf{B} = [\mathbf{A}_{.1} \otimes \mathbf{B}_{.1}, \mathbf{A}_{.2} \otimes \mathbf{B}_{.2}, \cdots, \mathbf{A}_{.J} \otimes \mathbf{B}_{.J}]$$

This is a matrix that is partitioned into J column-blocks, the j th block being equal to the Kronecker product of the j th column of $\mathbf{A}$ with the j th column of $\mathbf{B}$. It is said that $\mathbf{A} \diamond \mathbf{B}$ is a columnwise Kronecker product of $\mathbf{A}$ and $\mathbf{B}$.

Khatri-Rao product

Given $\mathbf{A} \in \mathbb{K}^{I \times J}$ and $\mathbf{B} \in \mathbb{K}^{K \times J}$ having the same number of columns, the Khatri-Rao product of $\mathbf{A}$ with $\mathbf{B}$, denoted by $\mathbf{A} \diamond \mathbf{B} \in \mathbb{K}^{IK \times J}$, is defined as:

$$\mathbf{A} \diamond \mathbf{B} = [\mathbf{A}_{.1} \otimes \mathbf{B}_{.1}, \mathbf{A}_{.2} \otimes \mathbf{B}_{.2}, \cdots, \mathbf{A}_{.J} \otimes \mathbf{B}_{.J}]$$

This is a matrix that is partitioned into J column-blocks, the j th block being equal to the Kronecker product of the j th column of $\mathbf{A}$ with the j th column of $\mathbf{B}$. It is said that $\mathbf{A} \diamond \mathbf{B}$ is a columnwise Kronecker product of $\mathbf{A}$ and $\mathbf{B}$.

定理

The Khatri-Rao product can also be written as a matrix partitioned into I row-blocks:

$$\mathbf{A} \diamond \mathbf{B} = \begin{bmatrix} \mathbf{B}\mathbf{D}_1(\mathbf{A}) \\ \mathbf{B}\mathbf{D}_2(\mathbf{A}) \\ \vdots \\ \mathbf{B}\mathbf{D}_I(\mathbf{A}) \end{bmatrix}$$

where $\mathbf{D}_i(\mathbf{A}) = \operatorname{diag}(a_{i1}, a_{i2}, \cdots, a_{iJ})$ refers to the diagonal matrix whose diagonal elements are the elements of the ith row of $\mathbf{A}$.

1 Introduction

2 Submatrices

3 Matrix products and partitioned matrices

4 **Special cases of partitioned matrices**

5 Determinants of partitioned matrices

Block-diagonal matrices

A square matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$ partitioned into (R, R) blocks, of diagonal blocks $\mathbf{A}_{rr} \in \mathbb{K}^{n_r \times n_r}$, with $r \in \langle R \rangle$ and $\sum_{r=1}^{R} n_r = n$, whose off-diagonal blocks are zero, is called a block-diagonal matrix and can be written as:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0}_{n_1 \times n_2} & \cdots & \mathbf{0}_{n_1 \times n_R} \\ \mathbf{0}_{n_2 \times n_1} & \mathbf{A}_{22} & \cdots & \mathbf{0}_{n_2 \times n_R} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n_R \times n_1} & \mathbf{0}_{n_R \times n_2} & \cdots & \boxed{\mathbf{A}_{RR}} \end{bmatrix}$$

blk diag

It is also written $\mathrm{diag}(\mathbf{A}_{11}, \mathbf{A}_{22}, \cdots, \mathbf{A}_{RR})$ or simply$\mathrm{diag}\,(\mathbf{A}_{rr})$ with the number R of diagonal blocks implied.

Signature matrices

The signature matrix of a symmetric matrix $\mathbf{A}$, of full rank, is a diagonal matrix whose diagonal elements are equal to 1 or -1:

$$\mathbf{S} = \mathrm{diag}(\underbrace{1, \cdots, 1}_{p \text{ terms}}, \underbrace{-1, \cdots, -1}_{q \text{ terms}}) \text{ with } p \geq 0, q \geq 0$$

p and q correspond to the numbers of positive and negative eigenvalues of $\mathbf{A}$, respectively. A signature matrix is thus a block-diagonal matrix consisting of two diagonal blocks $\mathbf{I}_p$ and $-\mathbf{I}_q$:

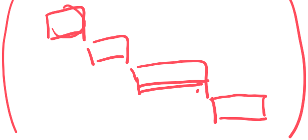
$$\mathbf{S} = \begin{bmatrix} \mathbf{I}_p & \mathbf{0}_{p \times q} \\ \mathbf{0}_{q \times p} & -\mathbf{I}_q \end{bmatrix}.$$

Direct sum

Given $\mathbf{A} \in \mathbb{K}^{I\times J}$ and $\mathbf{B} \in \mathbb{K}^{K\times L}$, the direct sum of $\mathbf{A}$ and $\mathbf{B}$, denoted by $\mathbf{A} \oplus \mathbf{B}$, is the block-diagonal matrix $\begin{bmatrix} \boxed{\mathbf{A}} & \cancel{\mathbf{0}_{I\times L}} \\ \cancel{\mathbf{0}_{K\times J}} & \boxed{\mathbf{B}} \end{bmatrix} \in \mathbb{K}^{(I+K)\times(J+L)}$. In the case of P matrices $\mathbf{A}^{(p)} \in \mathbb{K}^{I_p\times J_p}$, we have:

$$\bigoplus_{p=1} \mathbf{A}_p = \mathbf{A}_1 \oplus \mathbf{A}_2 \oplus \cdots \oplus \mathbf{A}_P \in \mathbb{K}^{\sum_{p=1}^{P} I_p \times \sum_{p=1}^{P} J_p}$$

$$= \operatorname{diag}(\mathbf{A}_1, \cdots, \mathbf{A}_P) = \operatorname{diag}(\mathbf{A}_p).$$



Jordan forms

A non-diagonalizable matrix $\mathbf{A} \in \mathbb{K}^{n\times n}$ can be transformed into a Jordan form:

$A^{100} \to P B^{100} P^{-1}$

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & & & \mathbf{0} \\ & \mathbf{B}_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{B}_p \end{bmatrix}, \tag{2}$$

① 求特征值

$(\lambda_i \underset{n\times n}{I} - A)x = 0$

$$\mathbf{B}_i = \begin{bmatrix} \lambda_i & 1 & & \mathbf{0} \\ & \lambda_i & 1 & \\ & & \ddots & 1 \\ \mathbf{0} & & & \lambda_i \end{bmatrix} \in \mathbb{K}^{n_i \times n_i},\ i \in \langle p \rangle,$$

where $\{\lambda_1, \cdots, \lambda_p\}$ are the eigenvalues of $\mathbf{A}$, and n_i is the multiplicity order of λ_i. Block-diagonal decomposition (2) into Jordan blocks $\mathbf{B}_i$, is called the Jordan form of A. This decomposition is little used in practice because its numerical determination may be unstable.

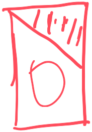
Block-triangular matrices

When $\mathbf{A}_{rs} = \mathbf{0}_{m_r, n_s}$ for $s < r$ in (1), that is:

$$\mathbf{A} = \begin{bmatrix} \boxed{\mathbf{A}_{11}} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1S} \\ \mathbf{0} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2S} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_{RS} \end{bmatrix},$$

A_{SS}

0 0 - - - 0



$\mathbf{A}$ is said to be an upper block-triangular matrix.

lower

Block-triangular matrices

When $\mathbf{A}_{rs} = \mathbf{0}_{m_r,n_s}$ for $s < r$ in (1), that is:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1S} \\ \mathbf{0} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_{RS} \end{bmatrix},$$

$\mathbf{A}$ is said to be an upper block-triangular matrix. Similarly, when $\mathbf{A}_{rs} = \mathbf{0}_{m_r,n_s}$ for $s > r$, that is:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{R1} & \mathbf{A}_{R2} & \cdots & \mathbf{A}_{RS} \end{bmatrix},$$

then $\mathbf{A}$ is called a lower block-triangular matrix.

Block Toeplitz and Hankel matrices

An $I \times J$ block-Toeplitz matrix is an $IM \times JN$ matrix partitioned in the form:

$$\mathbf{A} = \begin{bmatrix} \mathbf{F}_0 & \mathbf{F}_{-1} & \mathbf{F}_{-2} & \cdots & \mathbf{F}_{1-J} \\ \mathbf{F}_1 & \mathbf{F}_0 & \mathbf{F}_{-1} & \cdots & \mathbf{F}_{2-J} \\ \mathbf{F}_2 & \mathbf{F}_1 & \mathbf{F}_0 & \cdots & \mathbf{F}_{3-J} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{F}_{I-1} & \mathbf{F}_{I-2} & \mathbf{F}_{I-3} & \cdots & \mathbf{F}_0 \end{bmatrix}$$

where $\mathbf{F}_t \in \mathbb{C}^{M \times N}$, with $1 - J \leq t \leq I - 1$. When $M = N = 1$, we have a standard $I \times J$ Toeplitz matrix.

Block Toeplitz and Hankel matrices

An $I \times J$ block-Toeplitz matrix is an $IM \times JN$ matrix partitioned in the form:

$$\mathbf{A} = \begin{bmatrix} \mathbf{F}_0 & \mathbf{F}_{-1} & \mathbf{F}_{-2} & \cdots & \mathbf{F}_{1-J} \\ \mathbf{F}_1 & \mathbf{F}_0 & \mathbf{F}_{-1} & \cdots & \mathbf{F}_{2-J} \\ \mathbf{F}_2 & \mathbf{F}_1 & \mathbf{F}_0 & \cdots & \mathbf{F}_{3-J} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{F}_{I-1} & \mathbf{F}_{I-2} & \mathbf{F}_{I-3} & \cdots & \mathbf{F}_0 \end{bmatrix}$$

where $\mathbf{F}_t \in \mathbb{C}^{M \times N}$, with $1 - J \leq t \leq I - 1$. When $M = N = 1$, we have a standard $I \times J$ Toeplitz matrix. An $IJ \times IJ$ block-Hankel matrix is of the form:

$$\mathbf{A} = \begin{bmatrix} \mathbf{F}_0 & \mathbf{F}_1 & \mathbf{F}_2 & \cdots & \mathbf{F}_I \\ \mathbf{F}_1 & \mathbf{F}_2 & \mathbf{F}_3 & \cdots & \mathbf{F}_{I+1} \\ \mathbf{F}_2 & \mathbf{F}_3 & \mathbf{F}_4 & \cdots & \mathbf{F}_{I+2} \\ \vdots & \vdots & \vdots & . & \vdots \\ \mathbf{F}_I & \mathbf{F}_{I+1} & \mathbf{F}_{I+2} & \cdots & \mathbf{F}_{2I} \end{bmatrix}$$

where $\mathbf{F}_t$ is a $J \times J$ matrix for $t = 0, 1, \cdots, 2I$. As a Hankel matrix $\mathbf{A} = [a_{ij}] = [a_{i+j}]$, with $0 \leq i, j \leq I$, is determined by its first column and last row, a blockHankel matrix is such that $\mathbf{A} = [\mathbf{A}_{ij}] = [\mathbf{F}_{i+j}]$ with $0 \leq i, j \leq I$. When each block $\mathbf{F}_t$ is a Hankel matrix, then $\mathbf{A}$ is a block-Hankel matrix with Hankel blocks.

Transposition and conjugate transposition

The transposition (or conjugate transposition) of a matrix $\mathbf{A}$ partitioned into (R, S) blocks $\mathbf{A}_{rs}, r \in \langle R \rangle, s \in \langle S \rangle$, is obtained by transposing (or transconjugating) the blocks, followed by a blockwise transposition.

定理

For a matrix partitioned into $(2,2)$ blocks with square blocks of same dimensions:

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}.$$

$M^T = M$

$M^H = \begin{bmatrix} A^H & C^H \\ B^H & D^H \end{bmatrix}$

we have:

$$\mathbf{M} \text{ symmetric } \Leftrightarrow \mathbf{A}^T = \mathbf{A}, \mathbf{C} = \mathbf{B}^T, \mathbf{D}^T = \mathbf{D},$$

$$\mathbf{M} \text{ Hermitian } \Leftrightarrow \mathbf{A}^H = \mathbf{A}, \mathbf{C} = \mathbf{B}^H, \mathbf{D}^H = \mathbf{D},$$

that is, the diagonal blocks must be symmetric/Hermitian and the off-diagonal blocks transposed/conjugate transposed with respect to one another.

Trace

The trace of a partitioned matrix $\mathbf{A} = [\mathbf{A}_{rs}]$, with $r, s \in \langle R \rangle$, of dimensions (n, n) :

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1R} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2R} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{R1} & \mathbf{A}_{R2} & \cdots & \mathbf{A}_{RR} \end{bmatrix},$$

with $\dim(\mathbf{A}_{rr}) = (n_r, n_r)$ and $\sum_{r=1}^{R} n_r = n$, is given by:
$\mathrm{tr}(\mathbf{A}) = \sum_{r=1}^{R} \mathrm{tr}(\mathbf{A}_{rr})$.

Vectorization

Let us consider a balanced partitioning of $\mathbf{A}$ into (R, S) blocks of dimensions $P \times Q$. The partitioned matrix $\mathbf{A}$ can be vectorized column-blockwise (or rowblockwise), that is, by vectorizing each column (or row) of blocks, and then stacking the resulting vectors. The corresponding vectorization operators are denoted as $\text{vec}_c(.)$ and $\text{vec}_r(.)$, respectively.

Vectorization

Let us consider a balanced partitioning of $\mathbf{A}$ into (R, S) blocks of dimensions $P \times Q$. The partitioned matrix $\mathbf{A}$ can be vectorized column-blockwise (or rowblockwise), that is, by vectorizing each column (or row) of blocks, and then stacking the resulting vectors. The corresponding vectorization operators are denoted as $\text{vec}_c(.)$ and $\text{vec}_r(.)$, respectively.

例

For $R = S = 2$, we have:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \Rightarrow \text{vec}_c(\mathbf{A}) = \begin{bmatrix} \text{vec}\,(\mathbf{A}_{11}) \\ \text{vec}\,(\mathbf{A}_{21}) \\ \text{vec}\,(\mathbf{A}_{12}) \\ \text{vec}\,(\mathbf{A}_{22}) \end{bmatrix}, \text{vec}_r(\mathbf{A}) = \begin{bmatrix} \text{vec}\,(\mathbf{A}_{11}) \\ \text{vec}\,(\mathbf{A}_{12}) \\ \text{vec}\,(\mathbf{A}_{21}) \\ \text{vec}\,(\mathbf{A}_{22}) \end{bmatrix}.$$

Blockwise addition

Let $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{I \times J}$ be two matrices partitioned into blocks having the same dimensions $\mathbf{A}_{rs}, \mathbf{B}_{rs} \in \mathbb{K}^{I_r \times J_s}$, with $\sum_{r=1}^{R} I_r = I, \sum_{s=1}^{S} J_s = J$. Their sum is a partitioned matrix $\mathbf{C} = \mathbf{A} + \mathbf{B} = [\mathbf{A}_{rs} + \mathbf{B}_{rs}]$, with $\mathbf{C}_{rs} = \mathbf{A}_{rs} + \mathbf{B}_{rs} \in \mathbb{K}^{I_r \times J_s}$.

Blockwise addition

Let $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{I \times J}$ be two matrices partitioned into blocks having the same dimensions $\mathbf{A}_{rs}, \mathbf{B}_{rs} \in \mathbb{K}^{I_r \times J_s}$, with $\sum_{r=1}^{R} I_r = I, \sum_{s=1}^{S} J_s = J$. Their sum is a partitioned matrix $\mathbf{C} = \mathbf{A} + \mathbf{B} = [\mathbf{A}_{rs} + \mathbf{B}_{rs}]$, with
$\mathbf{C}_{rs} = \mathbf{A}_{rs} + \mathbf{B}_{rs} \in \mathbb{K}^{I_r \times J_s}$.

例

For $R = S = 2$, we have:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix},$$

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} \mathbf{A}_{11} + \mathbf{B}_{11} & \mathbf{A}_{12} + \mathbf{B}_{12} \\ \mathbf{A}_{21} + \mathbf{B}_{21} & \mathbf{A}_{22} + \mathbf{B}_{22} \end{bmatrix}$$

Blockwise multiplication

Let $\mathbf{A} \in \mathbb{K}^{I \times J}$ and $\mathbf{B} \in \mathbb{K}^{J \times L}$ be two matrices partitioned into blocks $\mathbf{A}_{rs} \in \mathbb{K}^{I_r \times J_s}$ and $\mathbf{B}_{sn} \in \mathbb{K}^{J_s \times L_n}$, with $\sum_{r=1}^{R} I_r = I, \sum_{s=1}^{S} J_s = J$, and $\sum_{n=1}^{N} L_n = L$. The product $\mathbf{C} = \mathbf{AB}$ is a matrix that is partitioned into blocks $\mathbf{C}_{rn} = \sum_{s=1}^{S} \mathbf{A}_{rs}\mathbf{B}_{sn} \in \mathbb{K}^{I_r \times L_n}$.

Blockwise multiplication

Let $\mathbf{A} \in \mathbb{K}^{I \times J}$ and $\mathbf{B} \in \mathbb{K}^{J \times L}$ be two matrices partitioned into blocks $\mathbf{A}_{rs} \in \mathbb{K}^{I_r \times J_s}$ and $\mathbf{B}_{sn} \in \mathbb{K}^{J_s \times L_n}$, with $\sum_{r=1}^{R} I_r = I, \sum_{s=1}^{S} J_s = J$, and $\sum_{n=1}^{N} L_n = L$. The product $\mathbf{C} = \mathbf{A}\mathbf{B}$ is a matrix that is partitioned into blocks $\mathbf{C}_{rn} = \sum_{s=1}^{S} \mathbf{A}_{rs}\mathbf{B}_{sn} \in \mathbb{K}^{I_r \times L_n}$.

例

For $R = S = N = 2$, we have:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}$$

$$\mathbf{A}\mathbf{B} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{bmatrix}.$$

Blockwise multiplication

Let $\mathbf{A} \in \mathbb{K}^{I \times J}$ and $\mathbf{B} \in \mathbb{K}^{J \times L}$ be two matrices partitioned into blocks $\mathbf{A}_{rs} \in \mathbb{K}^{I_r \times J_s}$ and $\mathbf{B}_{sn} \in \mathbb{K}^{J_s \times L_n}$, with $\sum_{r=1}^{R} I_r = I, \sum_{s=1}^{S} J_s = J$, and $\sum_{n=1}^{N} L_n = L$. The product $\mathbf{C} = \mathbf{AB}$ is a matrix that is partitioned into blocks $\mathbf{C}_{rn} = \sum_{s=1}^{S} \mathbf{A}_{rs}\mathbf{B}_{sn} \in \mathbb{K}^{I_r \times L_n}$.

例

For $R = S = N = 2$, we have:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{bmatrix}.$$

例

In the case of the product of two partitioned matrices built by adding a row and a column, respectively, we have:

$$\begin{bmatrix} \mathbf{X}^T \\ \mathbf{x}^T \end{bmatrix} [\mathbf{Y} \ \mathbf{y}] = \begin{bmatrix} \mathbf{X}^T\mathbf{Y} & \mathbf{X}^T\mathbf{y} \\ \mathbf{x}^T\mathbf{Y} & \mathbf{x}^T\mathbf{y} \end{bmatrix}.$$

$Yx^T + yx^T =$

Hadamard product of partitioned matrices

It should be remembered first that the Hadamard product of two matrices $\mathbf{A} \in \mathbb{K}^{I \times J}$ and $\mathbf{B} \in \mathbb{K}^{I \times J}$, of the same dimensions, gives a matrix $\mathbf{C} \in \mathbb{K}^{I \times J}$ defined as:

$$\mathbf{C} = \mathbf{A} \odot \mathbf{B} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1J}b_{1J} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2J}b_{2J} \\ \vdots & \vdots & \vdots & \\ a_{I1}b_{I1} & a_{I2}b_{I2} & \cdots & a_{IJ}b_{IJ} \end{bmatrix}$$

that is, $c_{ij} = a_{ij}b_{ij}$, and thus, $\mathbf{C} = [a_{ij}b_{ij}]$.

Hadamard product of partitioned matrices

It should be remembered first that the Hadamard product of two matrices $\mathbf{A} \in \mathbb{K}^{I \times J}$ and $\mathbf{B} \in \mathbb{K}^{I \times J}$, of the same dimensions, gives a matrix $\mathbf{C} \in \mathbb{K}^{I \times J}$ defined as:

$$\mathbf{C} = \mathbf{A} \odot \mathbf{B} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1J}b_{1J} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2J}b_{2J} \\ \vdots & \vdots & \vdots & \\ a_{I1}b_{I1} & a_{I2}b_{I2} & \cdots & a_{IJ}b_{IJ} \end{bmatrix}$$

that is, $c_{ij} = a_{ij}b_{ij}$, and thus, $\mathbf{C} = [a_{ij}b_{ij}]$.

Given $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{I \times J}$, partitioned into (R, S) blocks $\mathbf{A}_{rs}, \mathbf{B}_{rs} \in \mathbb{K}^{I_r \times J_s}$, with $I = \sum_{r=1}^{R} I_r$ and $J = \sum_{s=1}^{S} J_s$, then their Hadamard product $\mathbf{A} \odot \mathbf{B}$ is a partitioned matrix into (R, S) blocks $\mathbf{C}_{rs} = \mathbf{A}_{rs} \odot \mathbf{B}_{rs} \in \mathbb{K}^{I_r \times J_s}$, with $r \in \langle R \rangle, s \in \langle S \rangle$.

Hadamard product of partitioned matrices

It should be remembered first that the Hadamard product of two matrices $\mathbf{A} \in \mathbb{K}^{I \times J}$ and $\mathbf{B} \in \mathbb{K}^{I \times J}$, of the same dimensions, gives a matrix $\mathbf{C} \in \mathbb{K}^{I \times J}$ defined as:

$$\mathbf{C} = \mathbf{A} \odot \mathbf{B} = \begin{bmatrix} a_{11} b_{11} & a_{12} b_{12} & \cdots & a_{1J} b_{1J} \\ a_{21} b_{21} & a_{22} b_{22} & \cdots & a_{2J} b_{2J} \\ \vdots & \vdots & \vdots & \\ a_{I1} b_{I1} & a_{I2} b_{I2} & \cdots & a_{IJ} b_{IJ} \end{bmatrix}$$

that is, $c_{ij} = a_{ij} b_{ij}$, and thus, $\mathbf{C} = [a_{ij} b_{ij}]$.

Given $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{I \times J}$, partitioned into (R, S) blocks $\mathbf{A}_{rs}, \mathbf{B}_{rs} \in \mathbb{K}^{I_r \times J_s}$, with $I = \sum_{r=1}^{R} I_r$ and $J = \sum_{s=1}^{S} J_s$, then their Hadamard product $\mathbf{A} \odot \mathbf{B}$ is a partitioned matrix into (R, S) blocks $\mathbf{C}_{rs} = \mathbf{A}_{rs} \odot \mathbf{B}_{rs} \in \mathbb{K}^{I_r \times J_s}$, with $r \in \langle R \rangle, s \in \langle S \rangle$.

If all the blocks have the same dimensions $P \times Q$, that is, $I_r = P, \forall r \in \langle R \rangle$, and $J_s = Q, \forall s \in \langle S \rangle$, then $\mathbf{A}$ and $\mathbf{B}$, and consequently $\mathbf{A} \odot \mathbf{B}$ belong to the space denoted by $\mathcal{M}_{R \times S}(\mathcal{M}_{P \times Q}(\mathbb{K}))$.

Kronecker product of partitioned matrices

Given a matrix $\mathbf{A} \in \mathbb{K}^{I \times J}$ partitioned into (R, S) blocks $\mathbf{A}_{rs} \in \mathbb{K}^{I_r \times J_s}$, with $\sum_{r=1}^{R} I_r = I$ and $\sum_{s=1}^{S} J_s = J$, and a matrix $\mathbf{B} \in \mathbb{K}^{M \times N}$, then their Kronecker product $\mathbf{A} \otimes \mathbf{B}$ is a matrix partitioned into (R, S) blocks $\mathbf{A}_{rs} \otimes \mathbf{B} \in \mathbb{K}^{I_r M \times J_s N}$.

Kronecker product of partitioned matrices

Given a matrix $\mathbf{A} \in \mathbb{K}^{I \times J}$ partitioned into (R, S) blocks $\mathbf{A}_{rs} \in \mathbb{K}^{I_r \times J_s}$, with $\sum_{r=1}^{R} I_r = I$ and $\sum_{s=1}^{S} J_s = J$, and a matrix $\mathbf{B} \in \mathbb{K}^{M \times N}$, then their Kronecker product $\mathbf{A} \otimes \mathbf{B}$ is a matrix partitioned into (R, S) blocks $\mathbf{A}_{rs} \otimes \mathbf{B} \in \mathbb{K}^{I_r M \times J_s N}$.

例

For $R = 2, S = 3$, we have:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{bmatrix} \Rightarrow \mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} \mathbf{A}_{11} \otimes \mathbf{B} & \mathbf{A}_{12} \otimes \mathbf{B} & \mathbf{A}_{13} \otimes \mathbf{B} \\ \mathbf{A}_{21} \otimes \mathbf{B} & \mathbf{A}_{22} \otimes \mathbf{B} & \mathbf{A}_{23} \otimes \mathbf{B} \end{bmatrix}.$$

Kronecker product of partitioned matrices

More generally, in the case of two matrices partitioned into blocks $\mathbf{A}_{rs} \in \mathbb{K}^{I_r \times J_s}$, with $(r \in \langle R \rangle, s \in \langle S \rangle)$, and $\mathbf{B}_{mn} \in \mathbb{K}^{K_m \times L_n}$, with $(m \in \langle M \rangle, n \in \langle N \rangle)$, the block Kronecker product, called the Tracy-Singh product, is defined as:

$A_{12} \otimes (B_{mn})$

$$
\mathbf{A} \otimes_b \mathbf{B} = \begin{bmatrix}
\mathbf{A}_{11} \otimes \mathbf{B}_{11} & \cdots & \mathbf{A}_{11} \otimes \mathbf{B}_{1N} & \cdots & \mathbf{A}_{1S} \otimes \mathbf{B}_{11} & \cdots & \mathbf{A}_{1S} \otimes \mathbf{B}_{1N} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mathbf{A}_{11} \otimes \mathbf{B}_{M1} & \cdots & \mathbf{A}_{11} \otimes \mathbf{B}_{MN} & \cdots & \mathbf{A}_{1S} \otimes \mathbf{B}_{M1} & \cdots & \mathbf{A}_{1S} \otimes \mathbf{B}_{MN} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mathbf{A}_{R1} \otimes \mathbf{B}_{11} & \cdots & \mathbf{A}_{R1} \otimes \mathbf{B}_{1N} & \cdots & \mathbf{A}_{RS} \otimes \mathbf{B}_{11} & \cdots & \mathbf{A}_{RS} \otimes \mathbf{B}_{1N} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mathbf{A}_{R1} \otimes \mathbf{B}_{M1} & \cdots & \mathbf{A}_{R1} \otimes \mathbf{B}_{MN} & \cdots & \mathbf{A}_{RS} \otimes \mathbf{B}_{M1} & \cdots & \mathbf{A}_{RS} \otimes \mathbf{B}_{MN}
\end{bmatrix}
$$

of dimensions $\left(\sum_{r=1}^{R}\sum_{m=1}^{M} I_r K_m, \sum_{s=1}^{S}\sum_{n=1}^{N} J_s L_n\right)$. Note that if $R = S = M = N = 1$, this block Kronecker product becomes the classical Kronecker product, with $\mathbf{A} = \mathbf{A}_{11}$ and $\mathbf{B} = \mathbf{B}_{11}$.

Kronecker product of partitioned matrices

Another Kronecker product of partitioned matrices, called the strong Kronecker product and denoted by $|\otimes|$, was introduced for generating orthogonal matrices from Hadamard matrices. This Kronecker product is also used to represent tensor train decompositions. Given two matrices partitioned into blocks $\mathbf{A}_{rs} \in \mathbb{K}^{I \times J}$ and $\mathbf{B}_{sn} \in \mathbb{K}^{K \times L}$, with $r \in \langle R \rangle, s \in \langle S \rangle$, and $n \in \langle N \rangle$, the strong Kronecker product $\mathbf{A}| \otimes |\mathbf{B}$ is defined as the matrix partitioned into (R, N) blocks $\mathbf{C}_{rn} \in \mathbb{K}^{IK \times JL}$, with $r \in \langle R \rangle$ and $n \in \langle N \rangle$, such as:

$$\mathbf{C}_{rn} = \sum_{s=1}^{S} \mathbf{A}_{rs} \otimes \mathbf{B}_{sn}.$$

This operation, which is completely determined by the parameters (R, S, N), preserves the orthogonality.

- The matrix $\mathbf{1}_{RP\times SQ}$, whose all entries are equal to 1, is the identity element for $\odot$ in the space $\mathcal{M}_{R\times S}(\mathcal{M}_{P\times Q}(\mathbb{K}))$:

$$\mathbf{A}\odot\mathbf{1}_{RP\times SQ}=\mathbf{1}_{RP\times SQ}\odot\mathbf{A}=\mathbf{A},\forall\mathbf{A}\in\mathcal{M}_{R\times S}(\mathcal{M}_{P\times Q}(\mathbb{K}))$$

  Note that, for the Kronecker product, there is no identity element $\mathbf{E}$ such that $\mathbf{A}\otimes\mathbf{E}=\mathbf{E}\otimes\mathbf{A}=\mathbf{A},\forall\mathbf{A}$.

- Commutativity of $\odot$:

$$\mathbf{A}\odot\mathbf{B}=\mathbf{B}\odot\mathbf{A}$$

- Associativity of $\odot$ and $\otimes$:

$$\mathbf{A}\odot(\mathbf{B}\odot\mathbf{C})=(\mathbf{A}\odot\mathbf{B})\odot\mathbf{C}$$
$$\mathbf{A}\otimes(\mathbf{B}\otimes\mathbf{C})=(\mathbf{A}\otimes\mathbf{B})\otimes\mathbf{C}$$

- Distributivity of $\odot$ and $\otimes$ over the addition:

$$(\mathbf{A}+\mathbf{B})\odot\mathbf{C} = (\mathbf{A}\odot\mathbf{C}) + (\mathbf{B}\odot\mathbf{C})$$
$$\mathbf{A}\odot(\mathbf{B}+\mathbf{C}) = (\mathbf{A}\odot\mathbf{B}) + (\mathbf{A}\odot\mathbf{C})$$
$$(\mathbf{A}+\mathbf{B})\otimes\mathbf{C} = (\mathbf{A}\otimes\mathbf{C}) + (\mathbf{B}\otimes\mathbf{C})$$
$$\mathbf{A}\otimes(\mathbf{B}+\mathbf{C}) = (\mathbf{A}\otimes\mathbf{B}) + (\mathbf{A}\otimes\mathbf{C})$$

- Distributivity of $\odot$ and $\otimes$ over the scalar multiplication. For any $\lambda \in \mathbb{K}$:

$$\lambda(\mathbf{A}\odot\mathbf{B}) = (\lambda\mathbf{A})\odot\mathbf{B} = \mathbf{A}\odot(\lambda\mathbf{B})$$
$$\lambda(\mathbf{A}\otimes\mathbf{B}) = (\lambda\mathbf{A})\otimes\mathbf{B} = \mathbf{A}\otimes(\lambda\mathbf{B})$$

Elementary operations and elementary matrices

The elementary operations consist of:

- interchanging the ith and jth rows (columns):

$$\mathbf{A}_{i.} \leftrightarrow \mathbf{A}_{j.} \; (\mathbf{A}_{.i} \leftrightarrow \mathbf{A}_{.j})$$

- multiplying the elements of the ith row (column) by a scalar $k \neq 0$:

$$k\mathbf{A}_{i.} \rightarrow \mathbf{A}_{i.} \; (k\mathbf{A}_{.i} \rightarrow \mathbf{A}_{.i})$$

- adding to the elements of the ith row (column), the corresponding elements of the jth row (column) multiplied by k :

$$\mathbf{A}_{i.} + k\mathbf{A}_{j.} \rightarrow \mathbf{A}_{i.} \; (\mathbf{A}_{.i} + k\mathbf{A}_{.j} \rightarrow \mathbf{A}_{.i})$$

The three corresponding transformations, respectively, denoted by $\mathbf{P}_{ij}, \mathbf{P}_i(k)$, and $\mathbf{P}_{ij}(k)$ can be represented using the so-called elementary matrices.

Designating by $(p_1, \cdots, p_n)$ a permutation of $(1, \cdots, n)$, these elementary matrices are such that:

$$
\begin{aligned}
\mathbf{P}_{ij} &= [\mathbf{e}_{p_1} \cdots \mathbf{e}_{p_n}] \ \text{ with } \left[\begin{array}{c} \mathbf{e}_{p_i} = \mathbf{e}_j \text{ and } \mathbf{e}_{p_j} = \mathbf{e}_i \\ \mathbf{e}_{p_k} = \mathbf{e}_k \text{ if } k \neq i \text{ and } j \end{array} \right] \\
&= \mathbf{I} - (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T \\
\mathbf{P}_i(k) &= [\mathbf{e}_1, \cdots, \mathbf{e}_{i-1}, k\mathbf{e}_i, \mathbf{e}_{i+1}, \cdots, \mathbf{e}_n] \\
\mathbf{P}_{ij}(k) &= \left[\mathbf{e}_1, \cdots, \mathbf{e}_{j-1}, \mathbf{e}'_j, \mathbf{e}_{j+1}, \cdots, \mathbf{e}_n\right] \qquad \text{for rows} \\
&= \mathbf{I} + k\mathbf{e}_i\mathbf{e}_j^T \\
\mathbf{P}_{ij}(k) &= \left[\mathbf{e}_1, \cdots, \mathbf{e}_{i-1}, \mathbf{e}'_i, \mathbf{e}_{i+1}, \cdots, \mathbf{e}_n\right] \qquad \text{for columns} \\
&= \mathbf{I} + k\mathbf{e}_j\mathbf{e}_i^T
\end{aligned}
$$

$$\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}$$

where $\mathbf{e}'_j\,(\mathbf{e}'_i)$ is a vector consisting of 0s except its ith (jth) component equal to k and its jth (ith) component equal to 1.

例

Consider $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. For $\mathbf{P}_{12} = \mathbf{I}_3 - (\mathbf{e}_2 - \mathbf{e}_1)(\mathbf{e}_2 - \mathbf{e}_1)^T = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

we have:

$$\mathbf{P}_{12}\mathbf{A} = \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad \mathbf{A}\mathbf{P}_{12} = \begin{bmatrix} a_{12} & a_{11} & a_{13} \\ a_{22} & a_{21} & a_{23} \\ a_{32} & a_{31} & a_{33} \end{bmatrix}.$$

and

$$\mathbf{P}_{12}(k) = \mathbf{I}_3 + k\mathbf{e}_1\mathbf{e}_2^T = \begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \mathbf{P}_{12}(k)\mathbf{A} = \begin{bmatrix} a_{11} + ka_{21} & a_{12} + ka_{22} & a_{13} + ka_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\mathbf{P}_{12}(k) = \mathbf{I}_3 + k\mathbf{e}_2\mathbf{e}_1^T = \begin{bmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \mathbf{A}\mathbf{P}_{12}(k) = \begin{bmatrix} a_{11} + ka_{12} & a_{12} & a_{13} \\ a_{21} + ka_{22} & a_{22} & a_{23} \\ a_{31} + ka_{32} & a_{32} & a_{33} \end{bmatrix}.$$

Elementary operations and elementary matrices

Similarly, we can define elementary operations involving row-blocks or column-blocks of a partitioned matrix, to:

- interchange the i th and j th row-blocks (column-blocks);
- multiply the ith row-block (column-block) on the left-hand side (right-hand side) by a non-singular matrix;
- add the ith row-block (column-block) to the jth row-block (column-block) multiplied on the left-hand side (right-hand side) by a non-singular matrix.

Elementary operations and elementary matrices

Similarly, we can define elementary operations involving row-blocks or column-blocks of a partitioned matrix, to:

- interchange the i th and j th row-blocks (column-blocks);
- multiply the ith row-block (column-block) on the left-hand side (right-hand side) by a non-singular matrix;
- add the ith row-block (column-block) to the jth row-block (column-block) multiplied on the left-hand side (right-hand side) by a non-singular matrix.

For example, consider a matrix partitioned into $(2,2)$ blocks, with square diagonal blocks:

$$\mathbf{M} = \begin{bmatrix} \overset{n\times n}{\mathbf{A}} & \mathbf{B} \\ \mathbf{C} & \mathbf{D}_{m\times n} \end{bmatrix} \in \mathbb{C}^{(n+m)\times(n+m)}$$

$\mathbf{A}, \mathbf{E} \in \mathbb{C}^{n\times n}, \mathbf{D}, \mathbf{F} \in \mathbb{C}^{m\times m}, \mathbf{C}, \mathbf{G} \in \mathbb{C}^{m\times n}, \mathbf{B}, \mathbf{H} \in \mathbb{C}^{n\times m}$, with non-singular $\mathbf{E}$ and $\mathbf{F}$.

Elementary operations and elementary matrices

For the first type of elementary operation, we have:

$$\begin{bmatrix} \mathbf{0} & \mathbf{I}_m \\ \mathbf{I}_n & \mathbf{0} \end{bmatrix} \mathbf{M} = \begin{bmatrix} \mathbf{C} & \mathbf{D} \\ \mathbf{A} & \mathbf{B} \end{bmatrix}, \quad \mathbf{M} \begin{bmatrix} \mathbf{0} & \mathbf{I}_n \\ \mathbf{I}_m & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{B} & \mathbf{A} \\ \mathbf{D} & \mathbf{C} \end{bmatrix}.$$

Elementary operations and elementary matrices

For the first type of elementary operation, we have:

$$\begin{bmatrix} \mathbf{0} & \mathbf{I}_m \\ \mathbf{I}_n & \mathbf{0} \end{bmatrix} \mathrm{M} = \begin{bmatrix} \mathbf{C} & \mathbf{D} \\ \mathbf{A} & \mathbf{B} \end{bmatrix}, \quad \mathrm{M} \begin{bmatrix} \mathbf{0} & \mathbf{I}_n \\ \mathbf{I}_m & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{B} & \mathbf{A} \\ \mathbf{D} & \mathbf{C} \end{bmatrix}.$$

For the second type of elementary operation:

$$\begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} \mathrm{M} = \begin{bmatrix} \mathbf{EA} & \mathbf{EB} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \mathrm{M} \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} = \begin{bmatrix} \mathbf{AEB} & B \\ \mathbf{CE} & D \end{bmatrix},$$

$$\begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{F} \end{bmatrix} \mathrm{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{FC} & \mathbf{FD} \end{bmatrix}, \mathrm{M} \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{F} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{BF} \\ \mathbf{C} & \mathbf{DF} \end{bmatrix}.$$

For the third type of elementary operation:

$$\begin{bmatrix} \mathbf{I}_n & \mathbf{H} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} \mathbf{M} = \begin{bmatrix} \mathbf{A} + \mathbf{HC} & \mathbf{B} + \mathbf{HD} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$$

$$\mathbf{M} \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{G} & \mathbf{I}_m \end{bmatrix} = \begin{bmatrix} \mathbf{A} + \mathbf{BG} & \mathbf{B} \\ \mathbf{C} + \mathbf{DG} & \mathbf{D} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{G} & \mathbf{I}_m \end{bmatrix} \mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} + \mathbf{GA} & \mathbf{D} + \mathbf{GB} \end{bmatrix}$$

$$\mathbf{M} \begin{bmatrix} \mathbf{I}_n & \mathbf{H} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} + \mathbf{AH} \\ \mathbf{C} & \mathbf{D} + \mathbf{CH} \end{bmatrix}.$$

Inversion of partitioned matrices

This section is devoted to the inversion of 2×2 block matrices. We must note that for a 2×2 block matrix $\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$, whose blocks $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{D}$ have the dimensions $m \times n, m \times p, q \times n$ and $q \times p$, respectively, with $m + q = n + p$, its inverse $\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{bmatrix}$ is such that the blocks $\mathbf{E}, \mathbf{F}, \mathbf{G}$, and $\mathbf{H}$ must be of dimensions $n \times m$ $n \times q, p \times m$ and $p \times q$, respectively, in order to satisfy $\mathbf{M}\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{I}_m & \mathbf{0}_{m \times q} \\ \mathbf{0}_{q \times m} & \mathbf{I}_q \end{bmatrix}$ and $\mathbf{M}^{-1}\mathbf{M} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n \times p} \\ \mathbf{0}_{p \times n} & \mathbf{I}_p \end{bmatrix}$. So, we can conclude that the blocks of $\mathbf{M}^{-1}$ have the same dimensions as those of $\mathbf{M}^T$, and consequently, the partition of $\mathbf{M}^{-1}$ is transposed of that of $\mathbf{M}$.

Inversion of block-diagonal matrices

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix} \qquad M^{-1} = \begin{pmatrix} I & -A^{-1}B \\ & I \end{pmatrix} \begin{pmatrix} A^{-1} & \\ & D^{-1} \end{pmatrix}$$

$$M = \begin{bmatrix} I & B \\ 0 & I \end{bmatrix} \qquad M^{-1} = \begin{bmatrix} I & -B \\ 0 & I \end{bmatrix}$$

$$= \begin{pmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix}$$

Assuming that $\mathbf{A}$ and $\mathbf{D}$ are non-singular, we have:

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \quad \Rightarrow \quad \mathbf{M}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1} \end{bmatrix}.$$

For block upper and lower triangular matrices, with non-singular square diagonal blocks, we have:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{D}^{-1} \end{bmatrix}, \tag{3}$$

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{D}^{-1} \end{bmatrix}. \tag{4}$$

When $\mathbf{A} = \mathbf{I}_n$ and $\mathbf{D} = \mathbf{I}_m$, that is for unit block-triangular matrices, formulae (3) and (4) become:

$$\begin{bmatrix} \mathbf{I}_m & \mathbf{B} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I}_m & -\mathbf{B} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix},$$

$$\begin{bmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{C} & \mathbf{I}_n \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I}_m & \mathbf{0} \\ -\mathbf{C} & \mathbf{I}_n \end{bmatrix}.$$

Inversion of block-triangular matrices

$$\begin{pmatrix} A & B \\ C & 0 \end{pmatrix}^{-1} = \left[\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\begin{pmatrix} C & 0 \\ A & B \end{pmatrix}\right]^{-1} = \begin{pmatrix} C & 0 \\ A & B \end{pmatrix}^{-1}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}^{-1}$$

Similarly, with non-singular square off-diagonal blocks, we have:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} = \boxed{\begin{bmatrix} \mathbf{0} & \mathbf{C}^{-1} \\ \mathbf{B}^{-1} & -\mathbf{B}^{-1}\mathbf{A}\mathbf{C}^{-1} \end{bmatrix}},$$

$$\begin{bmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} -\mathbf{C}^{-1}\mathbf{D}\mathbf{B}^{-1} & \mathbf{C}^{-1} \\ \mathbf{B}^{-1} & \mathbf{0} \end{bmatrix},$$

with the following particular cases:

$$\begin{bmatrix} \mathbf{A} & \mathbf{I}_m \\ \mathbf{I}_n & \mathbf{0} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_n \\ \mathbf{I}_m & -\mathbf{A} \end{bmatrix}, \begin{bmatrix} \mathbf{0} & \mathbf{I}_m \\ \mathbf{I}_n & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} -\mathbf{D} & \mathbf{I}_n \\ \mathbf{I}_m & \mathbf{0} \end{bmatrix}.$$

When the partitioned matrix has no special structure, its inverse and its determinant are determined from block-triangular factorization, this factorization being itself obtained from block-diagonalization.

Block-triangularization and Schur complements

Assuming that $\mathbf{D}$ is non-singular, and applying the elementary transformation, the partitioned matrix $\mathbf{M}$, can be transformed into a lower block-triangular form. Indeed, by choosing $\mathbf{H} = -\mathbf{B}\mathbf{D}^{-1}$, we obtain:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

$$\begin{bmatrix} \mathbf{I}_n & -\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} \mathbf{M} = \begin{bmatrix} \mathbf{X}_{\mathbf{D}} & \mathbf{0} \\ \mathbf{C} & \mathbf{D} \end{bmatrix},$$

$$\mathbf{X}_{\mathbf{D}} = \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C},$$

where $\mathbf{X}_{\mathbf{D}}$, also denoted by $(\mathbf{M}/\mathbf{D})$, is called the Schur complement of $\mathbf{D}$ in $\mathbf{M}$. Similarly, assuming that $\mathbf{A}$ is non-singular and choosing $\mathbf{H} = -\mathbf{A}^{-1}\mathbf{B}$, we have:

$$\mathbf{M} \begin{bmatrix} \mathbf{I}_n - \mathbf{A}^{-1}\mathbf{B} & \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{C} & \mathbf{X}_{\mathbf{A}} \end{bmatrix},$$

$$\mathbf{X}_{\mathbf{A}} = \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B},$$

where $\mathbf{X}_{\mathbf{A}}$, also denoted by $(\mathbf{M}/\mathbf{A})$, is the Schur complement of $\mathbf{A}$ in $\mathbf{M}$. Similarly, elementary transformations as follows can be used to transform the partitioned matrix $\mathbf{M}$ into an upper block triangular form:

$$\begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_m \end{bmatrix} \mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{X}_{\mathbf{A}} \end{bmatrix}, \quad \begin{bmatrix} I & -BA^{-1} \\ & I \end{bmatrix} = \begin{pmatrix} A & 0 \\ 0 & X_A \end{pmatrix}$$

$$\mathbf{M} \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_m \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{\mathbf{D}} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}.$$

Block-diagonalization and block-factorization

Assuming that $\mathbf{A}$ is invertible, it is possible to put $\mathbf{M}$ in a block-diagonal form:

$$\begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_\mathbf{A} \end{bmatrix}. \tag{5}$$

-1 $\quad -1$ $\quad \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}$

Using the inversion formulae, the partitioned matrix $\mathbf{M}$ can be written in the following block-factorized form:

$\cancel{\begin{bmatrix} I & A^{-1}B \\ & I \end{bmatrix}} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cancel{\begin{bmatrix} I & \\ CA^{-1} & I \end{bmatrix}} = \begin{bmatrix} A^{-1} & \\ & X_A^{-1} \end{bmatrix} \quad \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix}$

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_\mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix}. \tag{6}$$

Similarly, assuming that $\mathbf{D}$ is invertible, a second block-diagonal form is obtained:

$$\begin{bmatrix} \mathbf{I}_n & -\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_m \end{bmatrix} = \begin{bmatrix} \mathbf{X}_\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}, \tag{7}$$

from which the following block-factorized form is deduced:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n & \mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{X}_\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_m \end{bmatrix}. \tag{8}$$

Block-inversion and partitioned inverse

The block-factorized form (6) gives the so-called Banachiewicz-Schur form, with $\mathbf{X_A} = \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$:

$$\begin{bmatrix} \overset{n\times n}{\mathbf{A}} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I}_n & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{X_A}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_m \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{X_A}^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{X_A}^{-1} \\ -\mathbf{X_A}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{X_A}^{-1} \end{bmatrix}. \qquad (9)$$

This inversion formula is valid if $\mathbf{A}$ and $\mathbf{X_A}$ are invertible.

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

$\overset{n\times n}{A} x = 0$

$\overset{m\times m}{X_A} y = 0$

Block-inversion and partitioned inverse

The block-factorized form (6) gives the so-called Banachiewicz-Schur form, with $\mathbf{X_A} = \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$:

$(A+UV^T)^{-1}$ $U, V \in \mathbb{R}^{n \times k}$

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I}_n & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{X_A}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_m \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{X_A}^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{X_A}^{-1} \\ -\mathbf{X_A}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{X_A}^{-1} \end{bmatrix}. \quad (9)$$

$(A+UV^T)^{-1}$

$(I+UV^T)^{-1}$

This inversion formula is valid if $\mathbf{A}$ and $\mathbf{X_A}$ are invertible. Similarly, the factorized form (8) leads to:

Sherman-Morrison-Woodbury

$$\begin{bmatrix} \mathbf{A} & \overset{U}{\mathbf{B}} \\ \underset{V^T}{\mathbf{C}} & \underset{I}{\mathbf{D}} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{X_D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & -\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{X_D}^{-1} & -\mathbf{X_D}^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{X_D}^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}\mathbf{X_D}^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix}. \quad (10)$$

This inversion formula is valid if $\mathbf{D}$ and $\mathbf{X_D} = \boxed{\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}}$ are invertible.

Other formulae for the partitioned 2×2 inverse

From (9) and (10), the following other forms of $\mathbf{M}^{-1}$ can be deduced:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{I}_m \end{bmatrix} \mathbf{X}_{\mathbf{A}}^{-1} \left[-\mathbf{C}\mathbf{A}^{-1} \mathbf{I}_m \right]$$
$$= \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1} \end{bmatrix} + \begin{bmatrix} \mathbf{I}_n \\ -\mathbf{D}^{-1}\mathbf{C} \end{bmatrix} \mathbf{X}_{\mathbf{D}}^{-1} \left[\mathbf{I}_n - \mathbf{B}\mathbf{D}^{-1} \right].$$

Combining formulae (9) and (10), we can also rewrite $\mathbf{M}^{-1}$ as:

$$\begin{aligned} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} &= \begin{bmatrix} \mathbf{X}_{\mathbf{D}}^{-1} & -\mathbf{X}_{\mathbf{D}}^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{X}_{\mathbf{A}}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{X}_{\mathbf{A}}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{X}_{\mathbf{D}}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{X}_{\mathbf{A}}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{X}_{\mathbf{D}}^{-1} & \mathbf{X}_{\mathbf{A}}^{-1} \end{bmatrix}, \end{aligned} \tag{11}$$

which gives the following block-factorizations:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{X}_{\mathbf{D}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{\mathbf{A}}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & -\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_m \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{I}_n & -\mathbf{A}^{-1}\mathbf{B} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{X}_{\mathbf{D}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{\mathbf{A}}^{-1} \end{bmatrix}.$$

Other formulae for the partitioned 2×2 inverse

By taking the first row-block of (9) and the second row-block of (10) , we get:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{X}_{\mathbf{A}}^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{X}_{\mathbf{A}}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{X}_{\mathbf{D}}^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}\mathbf{X}_{\mathbf{D}}^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix}.$$

Consider the following system of linear equations:

$$\left\{\begin{array}{ll} \mathbf{A}\mathbf{x}_1 + \mathbf{B}\mathbf{x}_2 & = \mathbf{y}_1 \\ \mathbf{C}\mathbf{x}_1 + \mathbf{D}\mathbf{x}_2 & = \mathbf{y}_2 \end{array}\right.,$$

or equivalently,

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}.$$

Applying the inversion formula (11) gives us the following solution:

$$\mathbf{x}_1 = \mathbf{X}_{\mathbf{D}}^{-1}\left(\mathbf{y}_1 - \mathbf{B}\mathbf{D}^{-1}\mathbf{y}_2\right),$$
$$\mathbf{x}_2 = \mathbf{X}_{\mathbf{A}}^{-1}\left(\mathbf{y}_2 - \mathbf{C}\mathbf{A}^{-1}\mathbf{y}_1\right),$$

where $\mathbf{X}_{\mathbf{A}}$ and $\mathbf{X}_{\mathbf{D}}$ are the Schur complements.

Inversion of a partitioned Gram matrix

Given a matrix partitioned into two column blocks $\mathbf{A} = [\mathbf{A}_1 \; \mathbf{A}_2]$, with $\mathbf{A}_1 \in \mathbb{R}^{n\times m}$ and $\mathbf{A}_2 \in \mathbb{R}^{n\times p}$, its Gram matrix $\mathbf{A}^T\mathbf{A}$ is partitioned as :

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} \mathbf{A}_1^T\mathbf{A}_1 & \mathbf{A}_1^T\mathbf{A}_2 \\ \mathbf{A}_2^T\mathbf{A}_1 & \mathbf{A}_2^T\mathbf{A}_2 \end{bmatrix}$$

Inversion of this type of matrix is employed for solving linear prediction problems or for estimating linear regression models. The application of inversion formula (11), with $\mathbf{A} = \mathbf{A}_1^T\mathbf{A}_1, \mathbf{B} = \mathbf{A}_1^T\mathbf{A}_2$, $\mathbf{C} = \mathbf{A}_2^T\mathbf{A}_1$, $\mathbf{D} = \mathbf{A}_2^T\mathbf{A}_2$, gives:

$$[\mathbf{A}^T\mathbf{A}]^{-1} = \begin{bmatrix} (\mathbf{A}_1^T\mathbf{P}_2^\perp\mathbf{A}_1)^{-1} & -(\mathbf{A}_1^T\mathbf{P}_2^\perp\mathbf{A}_1)^{-1}\mathbf{A}_1^T\mathbf{A}_2(\mathbf{A}_2^T\mathbf{A}_2)^{-1} \\ -(\mathbf{A}_2^T\mathbf{P}_1^\perp\mathbf{A}_2)^{-1}\mathbf{A}_2^T\mathbf{A}_1(\mathbf{A}_1^T\mathbf{A}_1)^{-1} & (\mathbf{A}_2^T\mathbf{P}_1^\perp\mathbf{A}_2)^{-1} \end{bmatrix}$$

where $\mathbf{P}_1^\perp$ and $\mathbf{P}_2^\perp$ are the orthogonal complements of orthogonal projection matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ on column spaces $C(\mathbf{A}_1)$ and $C(\mathbf{A}_2)$, respectively, that is:

$$\mathbf{P}_1^\perp = \mathbf{I}_n - \mathbf{P}_1 = \mathbf{I}_n - \mathbf{A}_1\left(\mathbf{A}_1^T\mathbf{A}_1\right)^{-1}\mathbf{A}_1^T,$$
$$\mathbf{P}_2^\perp = \mathbf{I}_n - \mathbf{P}_2 = \mathbf{I}_n - \mathbf{A}_2\left(\mathbf{A}_2^T\mathbf{A}_2\right)^{-1}\mathbf{A}_2^T.$$

Iterative inversion of a partitioned square matrix

$A^{-1} = (LU)^{-1} = U^{-1}L^{-1}$

$a_{11} = l_{11} \times u_{11}$

Consider the square matrix $\mathbf{M}_n$ of order n, partitioned into the following form:

$$\mathbf{M}_n = \begin{bmatrix} \mathbf{M}_{n-1} & \mathbf{c}_n \\ \mathbf{r}_n^T & \sigma_n \end{bmatrix}$$

where $\mathbf{M}_{n-1}$ is a square matrix of order $n-1$, and $\mathbf{c}_n, \mathbf{r}_n \in \mathbb{K}^{n-1}$. Assuming $\mathbf{M}_{n-1}$ and $\mathbf{M}_n$ are invertible, the application of the inversion formula (9) allows to perform the calculation of the inverse $\mathbf{M}_n^{-1}$ recursively with respect to order n, that is, in terms of $\mathbf{M}_{n-1}^{-1}$:

$$\mathbf{M}_n^{-1} = \begin{bmatrix} \mathbf{M}_{n-1}^{-1} + k_n \mathbf{M}_{n-1}^{-1}\mathbf{c}_n\mathbf{r}_n^T\mathbf{M}_{n-1}^{-1} & -k_n\mathbf{M}_{n-1}^{-1}\mathbf{c}_n \\ -k_n\mathbf{r}_n^T\mathbf{M}_{n-1}^{-1} & k_n \end{bmatrix},$$

where $k_n = (\sigma_n - \mathbf{r}_n^T\mathbf{M}_{n-1}^{-1}\mathbf{c}_n)^{-1}$ This recursive inversion formula will be used to demonstrate the Levinson-Durbin algorithm.

A^{-1} 或 $A=LU$

$\boxed{(A-BD^{-1}C)}\,x=b$

定理

Let $\mathbf{A} \in \mathbb{K}^{n\times n}, \mathbf{B} \in \mathbb{K}^{n\times m}, \mathbf{C} \in \mathbb{K}^{m\times n}$, *and* $\mathbf{D} \in \mathbb{K}^{m\times m}$. *By identifying the blocks* $(1,1)$ *of the right-hand sides of (9) and (10), it can be deduced that:*

$$\begin{aligned}\underbrace{\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1}}_{n\times n} &= \mathbf{A}^{-1}+\mathbf{A}^{-1}\mathbf{B}\mathbf{X}_{\mathbf{A}}^{-1}\mathbf{C}\mathbf{A}^{-1}\\ &= \mathbf{A}^{-1}+\mathbf{A}^{-1}\mathbf{B}\underbrace{\left(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right)^{-1}}_{m\times m}\mathbf{C}\mathbf{A}^{-1}\end{aligned} \tag{12}$$

This formula is known as the matrix inversion lemma. It is also called the Sherman-Morrison-Woodbury formula.

It should be noted that $\mathbf{X}_\mathbf{D} = \mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}$ is defined if $\mathbf{D}$ is invertible, and its inverse can be calculated using (12) if $\mathbf{A}$ and $\mathbf{X}_\mathbf{A} = \mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}$ are invertible.

Applications of the matrix inversion lemma

For different choices of $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$, the inversion lemma provides the following identities:

- For $\mathbf{D} = -\mathbf{I}_m$, the matrix inversion lemma (12) gives:

$$[\mathbf{A} + \mathbf{B}\mathbf{C}]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left[\mathbf{I}_m + \mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right]^{-1}\mathbf{C}\mathbf{A}^{-1}$$

- From this identity, it can be deduced that for $\mathbf{A} = \mathbf{I}_n$ and $\mathbf{C} = \mathbf{B}^H$

$$\left[\mathbf{I}_n + \mathbf{B}\mathbf{B}^H\right]^{-1} = \mathbf{I}_n - \mathbf{B}\left[\mathbf{I}_m + \mathbf{B}^H\mathbf{B}\right]^{-1}\mathbf{B}^H.$$

- For $m = n$, $\mathbf{D} = -\Delta^{-1}$ and $\mathbf{B} = \mathbf{C} = \mathbf{I}_n$

$$[\mathbf{A} + \Delta]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}[\mathbf{A}^{-1} + \Delta^{-1}]^{-1}\mathbf{A}^{-1}.$$

- For $m = 1$, $\mathbf{D} = -1/\alpha$, $\mathbf{B} = \mathbf{u} \in \mathbb{K}^n$, and $\mathbf{C}^T = \mathbf{v} \in \mathbb{K}^n$, assuming that $\alpha^{-1} + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u} \neq 0$, we get

$$\left[\mathbf{A} + \alpha\mathbf{u}\mathbf{v}^T\right]^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{\alpha^{-1} + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}$$

Generalized inverses of 2×2 block matrices

Consider the extension of the Banachiewicz-Schur form (9) to the case of singular or rectangular matrices partitioned into 2×2 blocks, more specifically with singular or rectangular submatrices, written as:

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \in \mathbb{K}^{(m+q)\times(n+p)}, \tag{13}$$

with $\mathbf{A} \in \mathbb{K}^{m\times n}, \mathbf{B} \in \mathbb{K}^{m\times p}, \mathbf{C} \in \mathbf{K}^{q\times n}$, and $\mathbf{D} \in \mathbf{K}^{q\times p}$. Unlike the case of a partitioned matrix $\mathbf{M}$ with square diagonal blocks, previously addressed, we now consider rectangular diagonal blocks, needing to define a generalized inverse for $\mathbf{M}$.

Generalized inverses of 2×2 block matrices

The notion of generalized inverse was introduced by Moore (1935), in a book published after his death. It was Penrose (1955) who demonstrated the uniqueness of the Moore generalized inverse, which explains the name given to the Moore-Penrose pseudo-inverse. This pseudo-inverse that generalizes the inverse of a regular square matrix to the case of rectangular matrices plays a very important role for solving systems of linear equations using the method of least squares.

Generalized inverses of 2×2 block matrices

The notion of generalized inverse was introduced by Moore (1935), in a book published after his death. It was Penrose (1955) who demonstrated the uniqueness of the Moore generalized inverse, which explains the name given to the Moore-Penrose pseudo-inverse. This pseudo-inverse that generalizes the inverse of a regular square matrix to the case of rectangular matrices plays a very important role for solving systems of linear equations using the method of least squares. Different types of generalized inverse $\mathbf{A}^{\#} \in \mathbb{K}^{n \times m}$ of a matrix $\mathbf{A} \in \mathbb{K}^{m \times n}$ can be defined according to the equations that are satisfied among:

1. $\mathbf{A}\mathbf{A}^{\#}\mathbf{A} = \mathbf{A}$
2. $\mathbf{A}^{\#}\mathbf{A}\mathbf{A}^{\#} = \mathbf{A}^{\#}$
3. $\left(\mathbf{A}\mathbf{A}^{\#}\right)^{H} = \mathbf{A}\mathbf{A}^{\#}$
4. $\left(\mathbf{A}^{\#}\mathbf{A}\right)^{H} = \mathbf{A}^{\#}\mathbf{A}$.

Note that (3) and (4) means that $\mathbf{A}\mathbf{A}^{\#}$ and $\mathbf{A}^{\#}\mathbf{A}$ are Hermitian, respectively, or symmetric in the real case.

Generalized inverses of 2×2 block matrices

Any inverse only satisfying conditions $\{c_1\}$, or $\{c_1, c_2\}$, or $\{c_1, c_2, c_3\}$, with $c_1, c_2, c_3 \in \{(1), (2), (3), (4)\}$, is denoted $\mathbf{A}^{\{c_1\}}, \mathbf{A}^{\{c_1,c_2\}}$, and $\mathbf{A}^{\{c_1,c_2,c_3\}}$, respectively.

- The properties of this type of inverse were studied by Ben-Israel and Greville (2001), where the inverses $\mathbf{A}^{\{1\}}, \mathbf{A}^{\{2\}}, \mathbf{A}^{\{1,2\}}$, and $\mathbf{A}^{(1,2,3)}$ are often called inner inverse, outer inverse, reflexive generalized inverse (or semi-inverse), and weak generalized inverse (or least-squares reflexive generalized inverse), respectively.
- For any matrix $\mathbf{A}$, there exists a unique matrix $\mathbf{A}^{\{1,2,3,4\}}$. This matrix corresponds to the Moore-Penrose pseudo-inverse of $\mathbf{A}$ and is often denoted by $\mathbf{A}^{\dagger}$.
- When $\mathbf{M}$ and $\mathbf{A}$ in (13) are singular, the Banachiewicz-Schur formula (9) can be extended by replacing the inverses of $\mathbf{A}$ and $\mathbf{X}_{\mathbf{A}}$ by generalized inverses $\mathbf{A}^{\#}$ and $\mathbf{X}_{\mathbf{A}}^{\#}$.

1 Introduction

2 Submatrices

3 Matrix products and partitioned matrices

4 Special cases of partitioned matrices

5 Determinants of partitioned matrices

Determinant of block-diagonal matrices

Assuming that $\mathbf{A}$ and $\mathbf{D}$ are square matrices, we have:

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \Rightarrow \det(\mathbf{M}) = \det(\mathbf{A})\det(\mathbf{D}).$$

Determinant of block-triangular matrices

For upper and lower block-triangular matrices, with square diagonal blocks, we have:

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \Rightarrow \det(\mathbf{M}) = \det(\mathbf{A})\det(\mathbf{D}).$$

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & 0 \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \Rightarrow \det(\mathbf{M}) = \det(\mathbf{A})\det(\mathbf{D}).$$

For unit block-triangular matrices $(\mathbf{A} = \mathbf{I}_n, \mathbf{D} = \mathbf{I}_m)$, we have $\det(\mathbf{M}) = 1$. More generally, we have:

$$\det \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1R} \\ & \mathbf{A}_{22} & & \vdots \\ & 0 & \ddots & \vdots \\ & & & \mathbf{A}_{RR} \end{bmatrix} = \prod_{r=1}^{R} \det\left(\mathbf{A}_{rr}\right).$$

From relations (5) and (7), the following expressions can be deduced:

$$\det \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \det(\mathbf{A})\det(\mathbf{X_A}) = \det(\mathbf{A})\det\left(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right) \quad (14)$$

$$= \det(\mathbf{D})\det(\mathbf{X_D}) = \det(\mathbf{D})\det\left(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right). \quad (15)$$

Determinants of specific partitioned matrices

$\left|\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\right| = 1\times4 - 3\times2$

定理

When $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ *are square matrices of order n, we have:*

2×2

- *If* $\mathbf{A}$ *and* $\mathbf{C}$ *commute:* $\det\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{bmatrix} = \det(\mathbf{AD} - \mathbf{CB})$.
- *If* $\mathbf{B}$ *and* $\mathbf{D}$ *commute:* $\det\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{bmatrix} = \det(\mathbf{DA} - \mathbf{BC})$.
- *If* $\mathbf{A}$ *and* $\mathbf{B}$ *commute:* $\det\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{bmatrix} = \det(\mathbf{DA} - \mathbf{CB})$.
- *If* $\mathbf{C}$ *and* $\mathbf{D}$ *commute:* $\det\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{bmatrix} = \det(\mathbf{AD} - \mathbf{BC})$.

Determinants of specific partitioned matrices

定理

The application of the previous formulae yields:

$$\det\begin{bmatrix} \mathbf{I}_n & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \det(\mathbf{D}-\mathbf{CB}); \qquad \det\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{I}_n \end{bmatrix} = \det(\mathbf{A}-\mathbf{BC}).$$

$$\det\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{A} & \mathbf{D} \end{bmatrix} = \det(\mathbf{A})\det(\mathbf{D}-\mathbf{B}); \det\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{B} \end{bmatrix} = \det(\mathbf{B})\det(\mathbf{A}-\mathbf{C}).$$

$$\det\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} = \det\begin{bmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = (-1)^n\det(\mathbf{B})\det(\mathbf{C}).$$

$$\begin{pmatrix} A & 0 \\ A & B \end{pmatrix}$$

Determinants of specific partitioned matrices

Defining:

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \text{ with } \mathbf{B}_{11} \in \mathbb{K}^{n\times n}, \mathbf{B}_{22} \in \mathbb{K}^{m\times m} \mathbf{A} \in \mathbb{K}^{n\times n}, \mathbf{C} \in \mathbb{K}^{m\times m},$$

and using the property $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$, as well as determinant formulae for block diagonal matrices and block-triangular matrices, the other following determinants can be deduced:

$$\det \begin{bmatrix} \mathbf{AB}_{11} & \mathbf{AB}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} = \det \left(\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix} \mathbf{B} \right) = \det(\mathbf{A})\det(\mathbf{B}),$$

and

$$\det \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} + \mathbf{CB}_{11} & \mathbf{B}_{22} + \mathbf{CB}_{12} \end{bmatrix} = \det \left(\begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{C} & \mathbf{I}_m \end{bmatrix} \mathbf{B} \right) = \det(\mathbf{B}).$$

Eigenvalues of CB and BC

定理

For $\mathbf{B} \in \mathbb{K}^{n\times m}$ *and* $\mathbf{C} \in \mathbb{K}^{m\times n}$, *we have:*

$$\lambda^n \det\left(\lambda \mathbf{I}_m - \mathbf{CB}\right) = \lambda^m \det\left(\lambda \mathbf{I}_n - \mathbf{BC}\right).$$

For $\lambda = 1$, *we obtain:*

$$\det\left(\mathbf{I}_m - \mathbf{CB}\right) = \det\left(\mathbf{I}_n - \mathbf{BC}\right).$$

Eigenvalues of CB and BC

定理

For $\mathbf{B} \in \mathbb{K}^{n \times m}$ *and* $\mathbf{C} \in \mathbb{K}^{m \times n}$*, we have:*

$$\lambda^n \det(\lambda \mathbf{I}_m - \mathbf{CB}) = \lambda^m \det(\lambda \mathbf{I}_n - \mathbf{BC}).$$

For $\lambda = 1$*, we obtain:*

$$\det(\mathbf{I}_m - \mathbf{CB}) = \det(\mathbf{I}_n - \mathbf{BC}).$$

证明.

By formulae (14) and (15), with $\mathbf{A} = \lambda \mathbf{I}_n$ and $\mathbf{D} = \mathbf{I}_m$, then:

$$\begin{aligned} \det \begin{bmatrix} \lambda \mathbf{I}_n & \mathbf{B} \\ \mathbf{C} & \mathbf{I}_m \end{bmatrix} &= \det(\lambda \mathbf{I}_n) \det\left(\mathbf{I}_m - \lambda^{-1} \mathbf{CB}\right) = \lambda^{n-m} \det(\lambda \mathbf{I}_m - \mathbf{CB}) \\ &= \det(\lambda \mathbf{I}_n - \mathbf{BC}) \end{aligned}$$

from which the identity (13) is deduced. $\square$

Rank of partitioned matrices

Let the partitioned matrix be:

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}.$$

Obviously, we have:

$$\begin{aligned} \mathrm{r}(\mathbf{M}) &\leq \mathrm{r}([\mathbf{A} \quad \mathbf{B}]) + \mathrm{r}([\mathbf{C} \quad \mathbf{D}]) \\ &\leq \mathrm{r}\left(\begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix}\right) + r\left(\begin{bmatrix} \mathbf{B} \\ \mathbf{D} \end{bmatrix}\right). \end{aligned}$$

In general, for a partitioned matrix $\mathbf{A} = [\mathbf{A}_{ij}]$, we have $\mathbf{r}(\mathbf{A}_{ij}) \leq \mathbf{r}(\mathbf{A})$
Based on block-factorization formulae (5) and (7), the following relations can be deduced, for $\mathbf{A} \in \mathbb{C}^{n\times n}$ and $\mathbf{D} \in \mathbb{C}^{m\times m}$:

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \Rightarrow \quad \begin{pmatrix} A & 0 \\ 0 & \boxed{X_A} \end{pmatrix}$$

$$\begin{aligned} \mathrm{r}(\mathbf{M}) = \mathrm{r}(\mathbf{A}) + \mathrm{r}(\mathbf{X_A}) &= \mathrm{r}(\mathbf{A}) + \mathrm{r}\left(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right) \\ = \mathrm{r}(\mathbf{D}) + \mathrm{r}(\mathbf{X_D}) &= \mathrm{r}(\mathbf{D}) + \mathrm{r}\left(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right). \end{aligned}$$

$Tx = b \quad T \in \mathbb{R}^{n\times n} \quad O(2n^3) \quad 10^6 \quad 10^{12}$

$O(4n^2) \quad 10^8$

Autoregressive (AR) processes or models are widely used in signal processing for the representation and classification of random signals. The autocorrelation function of such a process satisfies a system of linear equations in AR coefficients, called Yule–Walker equations. These equations form a Toeplitz system whose inversion can be achieved in a numerically efficient way, by means of the Levinson algorithm, also known as the Levinson-Durbin algorithm, which is an algorithm recursive with respect to the order of the model and plays a fundamental role in signal processing.

Yule–Walker equations

LU分解

$Tx = b$ T 正定 $-b$ $\frac{2}{3}n^3$

$T = LU$ ← 上三角

L: 单位下三角

$$\begin{pmatrix} t_0 & t_1 & \cdots & t_{n-1} \\ t_1 & \ddots & & \vdots \\ \vdots & & \ddots & t_1 \\ t_{n-1} & \cdots & t_1 & t_0 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = - \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_{n-1} \\ t_n \end{pmatrix}$$

t_n 任意

$\frac{1}{3}n^3$

$T = LL^T$ Cholesky 分解

L: 下三角

Yule-Walker

假设 $T_k x^{(k)} = -b^{(k)}$

T 的 K 阶顺序主子阵

$$T_k = \begin{pmatrix} t_0 & \cdots & t_{k-1} \\ t_1 & \ddots & \vdots \\ \vdots & & \\ t_{k-1} & \cdots & t_0 \end{pmatrix}$$

$$b^{(k)} = \begin{pmatrix} t_1 \\ \vdots \\ t_k \end{pmatrix}$$

$$E_k = \begin{pmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{pmatrix}$$

$T_{k+1} x^{(k+1)} = -b^{(k+1)}$

$$T_{k+1} = \begin{pmatrix} T_k & \begin{matrix} t_k \\ t_{k-1} \\ \vdots \end{matrix} \\ t_k \cdots & t_0 \end{pmatrix} = \begin{pmatrix} T_k & E_k b^{(k)} \\ (E_k b^{(k)})^T & t_0 \end{pmatrix}, \quad b^{(k+1)} = \begin{pmatrix} b^{(k)} \\ t_{k+1} \end{pmatrix}$$

Levinson-Durbin algorithm

$$x^{(k+1)} = \begin{pmatrix} z_k \\ \alpha \end{pmatrix} \qquad \begin{pmatrix} T_k & E_k b^{(k)} \\ (E_k b^{(k)})^T & t_0 \end{pmatrix} \begin{pmatrix} z_k \\ \alpha \end{pmatrix} = -\begin{pmatrix} b^{(k)} \\ t_{k+1} \end{pmatrix}$$

$$\begin{cases} T_k z_k + \alpha E_k b^{(k)} = -b^{(k)} & (1) \\ (E_k b^{(k)})^T z_k + t_0 \alpha = -t_{k+1} & (2) \end{cases} \qquad T_k E_k = E_k T_k$$

$$T_k z_k = -b^{(k)} - \alpha E_k b^{(k)} \qquad \boxed{T_k^{-1} E_k = E_k T_k^{-1}}$$

$$z_k = -T_k^{-1} b^{(k)} - \alpha T_k^{-1} E_k b^{(k)} \quad (3) \qquad \sum_{k=1}^{n} 2k \qquad n(n+1)$$

$$= x^{(k)} + \alpha E_k x^{(k)}$$