Block Matrix Formalization

Ethan Kalika

August 2, 2023

Abstract The point of this article is not to present novel discoveries but rather to make more available information that seems to be very scarcely documented. Block matrices and operations on them are defined in various sources, but very few sources document the proofs behind the workings of block matrices. These proofs are non-trivial and so I will detail here, a theoretical interpretation of block matrices that allows for the same functionality as that which is typically accepted and formalizes the language around the subject.

1 Definitions and Notation

1.1 Definition of a Block Matrix

The definition of a clock matrix is really abstract and will be developed over multiple subsections. Firstly, we must define two relations which will be used to partition a matrix.

1.1.1 The χ and ζ Relations

Let n and p be positive integers such that $n \geq p$, let $m_0 = 0$ and, let $\forall i \in \{1, \ldots, p\}, m_i \in \mathbb{Z}^+$. Now lets define two relations as follows: $\chi : \{1, \ldots, n\} \to \{1, \ldots, p\}$ and $\zeta : \{1, \ldots, n\} \to \{1, \ldots, \max\limits_{i \in \{1,\ldots,p\}} m_i\}$ where

$$\chi(x) = \min_{k \in \{1,\ldots,p\}} \{x - \sum_{i=0}^{k} m_i | x > \sum_{i=0}^{k} m_i\}$$

and

$$\zeta(x) = \min_{k \in \{1,\ldots,p\}} \{k | x < \sum_{i=0}^{k} m_i\}.$$

These two relations, for any $x \in \{1, \ldots, n\}$ define two integers $\chi(x)$ and $\zeta(x)$.

1.1.2 Theorem of $\chi - \zeta$ Bijectivity

Because χ and ζ are relations from $\{1, \ldots, n\}$ to $\{1, \ldots, p\}$ and from $\{1, \ldots, n\}$ to $\{1, \ldots, \max\limits_{i \in \{1,\ldots,p\}} m_i\}$ respectively, they are also relations from any subset of

$\{1, \ldots, n\}$ to $\{1, \ldots, p\}$ and $\{1, \ldots, \max_{i \in \{1,\ldots,p\}} m_i\}$ respectively.

Theorem $\forall i \in \{1, \ldots, p\}$, let $I_i = \{\sum_{y=0}^{i-1}(m_y) + 1, \ldots, \sum_{y=0}^{i} m_y\}$, then $\forall i \in \{1, \ldots, p\}, \chi$ is a bijection from I_i to $\{1, \ldots, m_i\} \wedge \zeta$ maps any $x \in I_i$ to i.

Proof For the duration of this proof let $i \in \{1, \ldots, p\}$, now we will show that $\forall i, \zeta$ maps any $x \in I_i$ to i. Let x be an element of I_i for some i, then by the definition of I_i, $\sum_{y=0}^{i-1} m_y < x \leq \sum_{y=0}^{i} m_y$. Then by the definition of ζ, $\zeta(x) = i$.

Now lets show that $\forall i, \chi$ is a bijection from I_i to $\{1, \ldots, m_i\}$. To do this we will first show that χ is a function between these sets, then that it is an injection, and finally that it is a surjection. To show that it is a function between the two sets we must show, firstly, that, statement $(*)$, $\forall i, (\forall x \in I_i, \exists z \in \{1, \ldots, m_i\} | (x, z) \in \chi$, is true, and secondly, that, statement $(**)$, $\forall i, (\forall z_1, z_2 \in \{1, \ldots, m_i\}, (\forall x \in I_i, ((x, z_1) \in \chi \wedge (x, z_2) \in \chi) \Rightarrow z_1 = z_2))$, is true.

First, lets show the former. Suppose that for some i, $x \in I_i$. Then $\sum_{y=0}^{i-1}(m_y) + 1 \leq x \leq \sum_{y=0}^{i} m_y$. Also $\chi(x) = \min_{k \in \{1,\ldots,q\}} \{x - \sum_{y=0}^{k} m_y | x > \sum_{y=0}^{k} m_y\}$. To minimize $x - \sum_{y=0}^{k} m_y$ we must find the maximum value of k such that $x > \sum_{y=0}^{k} m_y$, we will call this value k'. $k' < i$ because $x \leq \sum_{y=0}^{i} m_y$. Now, if we can show that $k = i - 1$ does not violate the condition, $x > \sum_{y=0}^{k} m_y$, then we know that $k' = i - 1$. $x \geq \sum_{y=0}^{i-1}(m_y) + 1 > \sum_{y=0}^{i-1} m_y$, therefor $k = i - 1$ satisfies the condition and $k' = i - 1$. So $x \in I_i \Rightarrow (\sum_{y=0}^{i-1}(m_y) + 1 \leq x \leq \sum_{y=0}^{i} m_y \wedge \chi(x) = x - \sum_{y=0}^{i-1} m_y)$. Then,

$$\begin{aligned}
x \in I_i &\Rightarrow \sum_{y=0}^{i-1}(m_y) + 1 \leq x \leq \sum_{y=0}^{i} m_y \\
&\Rightarrow 1 \leq x - \sum_{y=0}^{i-1} m_y \leq \sum_{y=0}^{i} m_y - \sum_{y=0}^{i-1} m_y \\
&\Rightarrow 1 \leq \chi(x) \leq m_i
\end{aligned}$$

That is if we let x be a p.b.a.c. element of I_i we find that $\chi(x)$ is between 1 and m_i. In other words we have shown that statement $(*)$ is true.

1.1.3 Partitioning Vectors

Given an n dimensional vector $\vec{v} = x_1\vec{e_1} + \cdots + x_n\vec{e_n}$ we can us these two functions to write a representation of $\vec{v}$ as an ordered tuple of vectors as follows:

$[x_1, \ldots, x_n] \rightarrow [A_1, \ldots, A_p]$ where $\forall i \in \{1, \ldots, p\}, A_i \in \mathbb{R}^{m_i}$ and x_i is the $\chi(i)^{th}$ entry of $A_{\zeta(i)}$. By the theorem of $\chi - \zeta$ bijectivity we know that $\zeta(i)$ takes on all values from 1 to p when i takes on all values from 1 to n, and that $\forall i \in \{1, \ldots, p\}$ the entries of A_i are defined by χ because χ is a bijection onto $\{1, \ldots, m_i\}$ which is the set of all integers from 1 to the length of A_i. Then we know that all entries of $[A_1, \ldots, A_p]$ can be defined in this way and $\forall i \in \{1, \ldots, n\}, x_i$ will be assigned a unique and position consisting of a respective A vector and position within A. That is, there is a one to one correspondence between the positions of $\vec{v}$ and $[A_1, \ldots, A_n]$. We will refer to the process of representing a vector in this way as partitioning a vector with respect to χ and ζ.

1.1.4 Partitioning Matrices and Defining the Block Matrix

Now lets consider an $m \times n$ matrix B. We can partition each row of this matrix with respect to χ and ζ which are functions defined in a way analogous to the one specified above and partition each column of B with respect to χ' and ζ' which are other functions defined in a way analogous to the method specified above. Then χ, ζ, χ', and ζ', define a one to one correspondence between the positions of B and position in partitionings of two vectors, the partitioning of the row and column of that position. We will denote a block matrix as a capital letter with a "$*$" in the superscript. The convention will also be that m and n with subscripts will be used to represent the number of rows and columns in the blocks of a block matrix respectively. There will be more about m and n notation in the anatomy section.

1.2 Anatomy of a Block Matrix Part 1

The ij^{th} block of a block matrix is the ij^{th} entry in the array of matrices and the ij^{th} entry of a block matrix is the ij^{th} entry of the matrix of the elements of the blocks. We will denote the ij^{th} block of a block matrix with the same capital letter as the block matrix with a subscript of ij. That is, if B^* is a block matrix then B_{ij} refers to ij^{th} block of the block matrix. We will denote the ij^{th} entry of a block matrix with the lowercase letter corresponding to the capital letter with a subscript ij. That is, if B^* is a block matrix then b_{ij} will denote the ij^{th} entry in B^*.

The i^{th} row of a block matrix, represented as R_i, is the i^{th} row of the matrix of entries in the blocks. The i^{th} column of a block matrix, which will be denoted C_i, is the i^{th} column of the matrix of the entries in the blocks. It goes without saying then, that if M^* is a block matrix with r rows and c columns then $\forall 1 \leq i \leq r, R_i \in \mathbb{R}^c$ and $\forall 1 \leq i \leq c, C_i \in \mathbb{R}^r$. It is important to note here that we assumed that M^* had c columns and r rowsThe i^{th} horizontal strip of a block matrix is a row of the $p \times q$ array of matrices. The i^{th} horizontal strip of a block matrix must have n_i rows by the definition of a block matrix, it must have as many columns as