Block Matrix Formalization

Ethan Kalika

September 3, 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 $\forall i \in \{1, \dots, p\}, m_i \in \mathbb{Z}^+ \mid \sum_{i=0}^{p} m_i = n$. Now lets define the following relation:$\chi : \{1, \dots, n\} \to \{1, \dots, p\}$ where

$$\chi(x) = \min_{k \in \{1, \dots, p\}} \{k \mid x \leq \sum_{i=0}^{k} m_i\}.$$

Lets also define the following relation on $\{1, \dots, n\}$:

$$\zeta(x) = x - \sum_{i=0}^{\chi(x)-1} m_i.$$

It is important to keep in mind that $\zeta : \{1, \dots, n\} \to \{1, \dots, m_{\chi(x)}\}$.

Proof. We prove this in three parts. Firstly, it is trivial that $\forall x \in \{1, \dots, n\}, \zeta(x) \in \mathbb{Z}$ because the relation involves adding and subtracting integers and integers are

closed under addition and subtraction. Secondly since $k = \chi(x)$ is the smallest value for which $x \leq \sum_{i=0}^{k}$, we know that $x > \sum_{i=0}^{\chi(x)-1} m_i$. Then, since $\zeta(x) = x - \sum_{i=0}^{\chi(x)-1} m_i$ we know that $\zeta(x) > 0$. Thirdly, we must show that $\forall x \in \{1, \ldots, n\}, \zeta(x) \leq m_{\chi(x)}$. We do this by contradiction. Suppose not, that is that, $\exists x \in \{1, \ldots, n\} \ni \zeta(x) > m_{\chi(x)}$. This would imply that:

$$x - \sum_{i=0}^{\chi(x)-1} m_i > m_{\chi(x)}$$

$$x > \sum_{i=0}^{\chi(x)} m_i$$

This is a contradiction because by the definition of $\chi(x)$, $k = \chi(x)$ is the smallest number for which $x \leq \sum_{i=0}^{k} m_i$, but this argument tells us that $x > \sum_{i=0}^{\chi(x)} m_i$. Then we know that $\forall x \in \{1, \ldots, n\}, 0 < \zeta(x) \leq m_{\chi(x)} \wedge \zeta(x) \in \mathbb{Z}$. Therefor, $\forall x \in \{1, \ldots, n\}, \zeta(x) \in \{1, \ldots, m_{\chi(x)}\}$. $\qquad\square$

The two relations χ and ζ, for any $x \in \{1, \ldots, n\}$ define two integers $\chi(x)$ and $\zeta(x)$. This idea will be very important in the sections to come.

1.1.2   Properties of χ and ζ

Because χ and ζ are relations from $\{1, \ldots, n\}$ to $\{1, \ldots, p\}$ and $\{1, \ldots, m_{\chi(x)}\}$ respectively, they are also relations from any subset of $\{1, \ldots, n\}$ to $\{1, \ldots, p\}$ and $\{1, \ldots, m_{\chi(x)}\}$ respectively. Lets also pose the following definition: $\forall i \in \{1, \ldots, p\}, I_i = \{\sum_{y=0}^{i-1}(m_y) + 1, \ldots, \sum_{y=0}^{i} m_y\}$, that is let $\forall i \in \{1, \ldots, p\}, I_i$ be the set of consecutive integers from $\sum_{y=0}^{i-1}(m_y) + 1$ to $\sum_{y=0}^{i} m_y$ inclusive.

χ - ζ **Bijectivity Theorem** $\forall i \in \{1, \ldots, p\}, \zeta$ is a bijection from I_i to $\{1, \ldots, m_i\} \wedge \chi$ maps any $x \in I_i$ to i.

Proof. For the duration of this proof let i be a p.b.a.c. element of $\{1, \ldots, p\}$, now we will show that $\forall i, \chi$ 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, $\sum_{y=0}^{i-1}(m_y) + 1 \leq x \leq \sum_{y=0}^{i} m_y$. Then by the definition of χ, $\chi(x) = i$.

Now lets show that $\forall i, \zeta$ is a bijection from I_i to $\{1, \ldots, m_i\}$. To do this we will first show that ζ 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 $(*)$, $\forall i, (\forall x \in I_i, \exists z \in \{1, \ldots, m_i\} \ni (x, z) \in \zeta$, is true, and secondly, that, statement $(**)$, $\forall i, (\forall z_1, z_2 \in \{1, \ldots, m_i\}, (\forall x \in I_i, ((x, z_1) \in \zeta \wedge (x, z_2) \in \zeta) \Rightarrow z_1 = z_2))$, is true.

First, lets show $(*)$. 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 $\chi(x) = \min_{k \in \{1,\dots,q\}} \{x - \sum_{y=0}^{k} m_y | x > \sum_{y=0}^{k} m_y\}$. 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 $x \le \sum_{y=0}^{i} m_y$. 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 $k' = i - 1$. $x \ge \sum_{y=0}^{i-1}(m_y) + 1 > \sum_{y=0}^{i-1} m_y$, therefor $k = i - 1$ satisfies the condition and $k' = i - 1$. So $x \in I_i \Rightarrow (\sum_{y=0}^{i-1}(m_y) + 1 \le x \le \sum_{y=0}^{i} m_y \wedge \chi(x) = x - \sum_{y=0}^{i-1} m_y)$. Then,

$$x \in I_i \Rightarrow \sum_{y=0}^{i-1}(m_y) + 1 \le x \le \sum_{y=0}^{i} m_y$$

$$\Rightarrow 1 \le x - \sum_{y=0}^{i-1} m_y \le \sum_{y=0}^{i} m_y - \sum_{y=0}^{i-1} m_y$$

$$\Rightarrow 1 \le \chi(x) \le m_i$$

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.

$\square$

I-Partition Theorem $\{I_1, \dots, I_p\}$ is a partitioning of the set $\{1, \dots, n\}$.

Proof. First lets show that $\forall a \in \{1, \dots, p\}(\forall b \in \{1, \dots, p\}, a \ne b \to I_a \cap I_b = \emptyset)$. Lets suppose WLOG that $a > b$. Then by the definition of I_i, $\forall z \in I_b, z \le \sum_{y=0}^{b} m_y$ and $\forall x \in I_a, x \ge \sum_{y=0}^{a-1}(m_y) + 1$. Since a and b are integers and $a > b$ we know that $a - 1 \ge b$, then, because by definition $m_y > 0$ we know that $\sum_{y=0}^{a-1} m_y \ge \sum_{y=0}^{b} m_y$. Then, $\forall x \in I_a, (\forall z \in I_b, x \ge \sum_{y=0}^{a-1}(m_y) + 1 > \sum_{y=0}^{a-1} m_y \ge \sum_{y=0}^{b} m_y \ge z)$, that is $\forall x \in I_a, (\forall z \in I_b, x > z)$, and so not two distinct sets I_i have any common elements.

Second lets show that $I_1 \cup, \dots, \cup I_p \subseteq \{1, \dots, n\}$. Let x be a p.b.a.c. element of $I_1 \cup, \dots, \cup I_p$, then $\sum_{y=0}^{0}(m_y) + 1 = 1 \le x \le \sum_{y=0}^{p} m_y = n$. That is $x \in I_1 \cup, \dots, \cup I_p \Rightarrow x \in \{1, \dots, n\}$.

Third, lets show that $\{1, \dots, n\} \subseteq I_1 \cup, \dots, \cup I_p$. We do this by proof by contradiction and division into cases. Suppose not, that is suppose $\exists x \in \{1, \dots, n\} \ni x \notin I_1 \cup, \dots, \cup I_n$, then we have three cases:

Case 1: $x < \min I_1$
Then $x < 1$. We also know that $x \in \{1, \dots, n\} \Rightarrow x \ge 1$. Then we have a contradiction.

Case 2: $\exists i \in \{2, \dots, p\} \ni \max I_{i-1} < x < \min I_i$.
Then, by lemma 2.1, $\max I_{i-1} < x < \max I_{i-1} + 1$ and $0 < x - \max I_{i-1} < 1$.

Since integers are closed under subtraction and $\max I_{i-1}$ and x are integers, this would mean that there is an integer between 0 and 1 which is a contradiction.

Case 3: $x > \max I_p$
By the definition of I_i we know that $\max I_p = \sum_{y=0}^{p} m_y = n$. Then we have a contradiction because $x \in \{1, \ldots, n\} \Rightarrow x \leq n$.

Then $\{I_1, \ldots, I_p\}$ is a partitioning of the set $\{1, \ldots, n\}$. $\qquad\square$

Corollary 2.1 χ is a surjection from $\{1, \ldots, n\}$ to $\{1, \ldots, p\}$.

Proof. We must show that $\forall y \in \{1, \ldots, p\}, \exists x \in \{1, \ldots, n\} \ni \chi(x) = y$. By I-Partition Theorem we know that $\{I_1, \ldots, I_p\}$ is a partition of $\{1, \ldots, n\}$, that is, that $\forall i \in \{1, \ldots, p\}, I_i \subset \{1, \ldots, n\}$. By the χ - ζ Bijectivity Theorem we know that χ maps any element of I_i to i. Since $\{1, \ldots, n\}$ contains all the elements in $I_1, I_2, \ldots$, and I_p it must contain at least one element, x, for which $\chi(x) = 1$, $\chi(x) = 2$, $\ldots$, and $\chi(x) = p$. $\qquad\square$

Lemma 2.1 $\forall i \in \{2, \ldots, p\}, \min I_i = \max I_{i-1} + 1$

Proof. By the definition of I_i, $\min I_i = \sum_{y=0}^{i-1}(m_y) + 1$ and $\max I_{i-1} = \sum_{y=0}^{i-1} m_y$.
$\qquad\square$

The purpose of these theorems to build on the conclusion from the last section by showing that $\forall x \in \{1, \ldots, n\}, (\chi(x), \zeta(x))$ is a unique ordered pair.

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 $\zeta(i)^{th}$ entry of $A_{\chi(i)}$. By corollary 1 we know that $\chi(i)$ takes on all values from 1 to p when i takes on all values from 1 to n, so all vectors in the set $\{A_1, \ldots, A_p\}$ are defined. By the Theorem of χ - ζ bijectivity we know that for every vector A_i all of its entries are defined by ζ because ζ is a bijection mapping to $\{1, \ldots, m_i\}$ which is the set of all integers from 1 to the length of A_i. Then we know that all positions in $[A_1, \ldots, A_p]$ 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 the positions of $[A_1, \ldots, A_n]$. 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 ith row of a block matrix, represented as R_i, is the ith row of the matrix of entries in the blocks. The ith column of a block matrix, which will be denoted C_i, is the ith 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 $\forall 1 \le i \le r, R_i \in \mathbb{R}^c$ and $\forall 1 \le i \le c, C_i \in \mathbb{R}^r$. 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 ith 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