8-27-2016

Invertibility of Submatrices of the Pascal Matrix and Birkhoff Interpolation

Scott N. Kersey

Georgia Southern University, skersey@georgiasouthern.edu

Follow this and additional works at: https://digitalcommons.georgiasouthern.edu/math-sci-facpubs

Part of the Mathematics Commons

Recommended Citation

https://digitalcommons.georgiasouthern.edu/math-sci-facpubs/532

This article is brought to you for free and open access by the Mathematical Sciences, Department of at Digital Commons@Georgia Southern. It has been accepted for inclusion in Mathematical Sciences Faculty Publications by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.
INVERTIBILITY OF SUBMATRICES OF THE PASCAL MATRIX AND
BIRKHOFF INTERPOLATION

SCOTT KERSEY

Abstract. The infinite upper triangular Pascal matrix is \( T = \binom{j}{i} \) for \( 0 \leq i, j \). It is easy
to see that any leading principle square submatrix is triangular with
determinant 1, hence invertible. In this paper, we investigate the invertibility of arbitrary square submatrices
\( T_{r,c} \) comprised of rows \( r = [r_0, \ldots, r_m] \) and columns \( c = [c_0, \ldots, c_m] \) of \( T \). We show that \( T_{r,c} \) is
invertible iff \( r \leq c \) (i.e., \( r_i \leq c_i \) for \( i = 0, \ldots, m \)), or equivalently, iff all diagonal entries are
nonzero. To prove this result we establish a connection between the invertibility of these
submatrices and polynomial interpolation. In particular, we apply the theory of Birkhoff
interpolation and Pólya systems.

Paper appears in: Journal of Mathematical Sciences: Advances and Applications, 41,
45–56 (2016).

1. Introduction

The infinite and order \( n + 1 \) upper triangular Pascal matrices are

\[
T := \begin{bmatrix}
1 & 1 & 1 & 1 & \cdots \\
0 & 1 & 2 & 3 & \cdots \\
0 & 0 & 1 & 3 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}
\quad \text{and} \quad
T_n := \begin{bmatrix}
1 & 1 & 1 & \cdots & \binom{n}{0} \\
0 & 1 & 2 & \cdots & \binom{n}{1} \\
0 & 0 & 1 & \cdots & \binom{n}{2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \binom{n}{n}
\end{bmatrix},
\]

for \( i, j = 0, 1, \ldots, \), with \( \binom{i}{j} := 0 \) if \( i > j \). Note that \( T_n \) is a finite square submatrix of
\( T \) determined by the first \( n + 1 \) rows and columns. Other (square) submatrices of \( T \) are
determined by selecting sequences of the rows and columns of \( T \). These submatrices can be
represented as

\[
T_{r,c} = \begin{bmatrix}
\binom{c_0}{r_0} & \binom{c_1}{r_0} & \cdots & \binom{c_m}{r_0} \\
\binom{c_0}{r_1} & \binom{c_1}{r_1} & \cdots & \binom{c_m}{r_1} \\
\vdots & \vdots & \ddots & \vdots \\
\binom{c_0}{r_m} & \binom{c_1}{r_m} & \cdots & \binom{c_m}{r_m}
\end{bmatrix}
\]

for some selections \( r = [r_0, \ldots, r_m] \) and \( c = [c_0, \ldots, c_m] \) of the rows and columns of \( T \),
respectively.

The main issue considered in this paper is invertibility of these submatrices. It is trivial
to see that \( T_n \) is invertible, with determinant 1. But what about other submatrices? For
example,
\[
T_{[0,1,2],[1,2,5]} = \begin{bmatrix}
1 & 1 & 1 \\
1 & 2 & 5 \\
0 & 1 & 10
\end{bmatrix}
\quad \text{and} \quad
T_{[1,3,4],[1,2,5]} = \begin{bmatrix}
1 & 2 & 5 \\
0 & 0 & 10 \\
0 & 0 & 5
\end{bmatrix}.
\]

The second matrix, a triangular matrix with a zero on the main diagonal, clearly has a zero determinant, hence is not invertible. However, while the first matrix is invertible, it is not obvious to see. In this paper, we will prove the following necessary and sufficient conditions for invertibility of these submatrices:

**Theorem 1.1.** Submatrices \( T_{r,c} \) of the upper Pascal matrix are invertible iff the following equivalent conditions hold:

- \( r \leq c \) (i.e., \( r_i \leq c_i \) for all \( i \)).
- There is no zero diagonal entry.

To our knowledge, this result has not appeared in the literature, other than in a preliminary draft of this paper in [4]. The motivation for this work comes from our work on dual basis in subspaces which first appeared in [3]. In that paper we use the theorem to demonstrate the linear independence of certain dual basis functions.

As it turns out, the problem of invertibility of submatrices of the Pascal matrix is connected to the theory of Birkhoff interpolation. This goes back to a paper of G. Pólya in [7], as used by J.M. Whittaker in [8], and later generalized by D. Ferguson in [2]. The more general Birkhoff interpolation problem (with more than two interpolation points) was presented by G. D. Birkhoff in [1]. In our work, we show that such submatrices of the Pascal matrix are invertible iff a certain 2-point Birkhoff interpolation problem satisfies the Pólya condition.

More recently, in [5] and [6], Birkhoff interpolation is generalized to lacunary interpolation, which could be applied to the problem investigated in this paper. While some of the techniques in their work are similar to ours, we were not aware of their work during the preparation of this paper, and so our results were obtained independently. Finally, while the work in this paper originates to [4], much of the original material has been removed and some content has been changed and/or corrected. In particular, in the original draft we refer to the submatrices as truncations of the Pascal matrix.

The remainder of this paper is organized as follows:

- In section (2), we show that our problem is equivalent to a certain two-point polynomial interpolation problem.
- In Section (3), we summarize pertinent aspects of Birkhoff interpolation and Pólya systems.
- In section (4), we prove Theorem 1.1.

2. The Pascal Matrix and Polynomial Interpolation

We begin this paper by establishing a connection between (submatrices of) the Pascal matrix and polynomial interpolation. Let \( \Lambda_\alpha \) be a row vector
\[
\Lambda_\alpha = \begin{bmatrix}
\delta_\alpha \\
\delta_\alpha D \\
\vdots \\
\delta_\alpha D^i \\
\vdots
\end{bmatrix}
\]
comprised of the functionals
\[ \delta_\alpha D^i : f \mapsto f^{(i)}(\alpha), \]
and let \( V \) be the sequence of monomials
\[ V = [1, (\cdot), (\cdot)^2, \ldots]. \]
Note that \( \Lambda^T_\alpha V \) is an infinite matrix. We will be interested in the maps \( \Lambda_0 \) and \( \Lambda_1 \), when \( \alpha = 0 \) and \( \alpha = 1 \), respectively. In particular, we have the following:

**Proposition 2.1.** The Pascal matrix \( T = [\binom{i}{j}] \) is equivalent to
\[ \Lambda^T_1 V = [\binom{\delta_1 D^i x^j}{i!}]. \]

**Proof.** Let \( A := \Lambda^T_1 V \). We will show that \( A = T \). For \( j \geq i \),
\[ A(i, j) = \frac{\delta_1 D^i}{i!} x^j = \frac{1}{i!} \frac{j!}{(j-i)!} x^{j-i} \bigg|_{x=1} = \binom{j}{i}. \]
For \( j < i \),
\[ A(i, j) = \frac{\delta_1 D^i}{i!} x^j = \frac{\delta_1}{i!} 0 = 0. \]
Hence, \( A = T \).

We can state a similar result for submatrices of the Pascal matrix. Let \( r = [r_0, \ldots, r_m] \) and \( c = [c_0, \ldots, c_m] \) be selections of the rows and and columns of \( T \), respectively. Let
\[ \Lambda_{\alpha, r} = [\binom{\delta_\alpha D^r}{r_0!}, \ldots, \binom{\delta_\alpha D^r}{r_m!}], \]
and
\[ V_c = [\binom{\cdot}{c_0}, \ldots, \binom{\cdot}{c_m}]. \]
Then, by restricting the the result in Proposition 2.1 to the row \( r \) and columns \( c \), we obtain the following:

**Proposition 2.2.** Let \( r = [r_0, \ldots, r_m] \) and \( c = [c_0, \ldots, c_m] \) be selections of the rows and columns of \( T \). Then,
\[ T_{r,c} = \Lambda^T_{1,r} V_c. \]

Hence, to establish the invertibility of \( T_{r,c} \), we can analyze \( \Lambda^T_{1,r} V_c \). To do so, we will establish a connection with two-point polynomial interpolation, which we do by augmenting the point functionals at \( x = 1 \) with functionals at \( x = 0 \). Let \( \bar{c} \) be the complement of \( c \) in \([0 : n]\), with \( n := \max\{r_m, c_m\} \). Let
\[ \Lambda_{r, \bar{c}} := [\Lambda_{1,r}, \Lambda_0, \bar{c}], \]
and
\[ V_{c, \bar{c}} = [V_c, V_\bar{c}] = [(\cdot)^{c_0}, \ldots, (\cdot)^{c_d}, (\cdot)^{r_0}, \ldots, (\cdot)^{r_{n-m}}]. \]
Note that \( V_{c, \bar{c}} \) is the power basis of degree \( n \) with the terms rearranged.

**Theorem 2.3.** \( T_{r,c} \) is invertible iff \( \Lambda^{T}_{1,r} V_{c, \bar{c}} \) is invertible.
Proof. The interpolation matrix can be expressed in block form as follows:

$$\Lambda_{r, \overline{c}}^T V_{c, \overline{c}} = \left[ \Lambda_{1, r}, \Lambda_{0, \overline{c}} \right]^T [V_c, V_{\overline{c}}] = \left[ \begin{array}{l|l}
\Lambda_{1, r}^T V_c & \Lambda_{1, r}^T V_{\overline{c}} \\
\Lambda_{0, c}^T V_c & \Lambda_{0, \overline{c}}^T V_{\overline{c}} \end{array} \right].$$

By Proposition $[2, 2]$ the upper left block is $T_{r, c}$. The lower left and lower right blocks comprises terms of the form $\frac{1}{n!} \delta_0 D^i = \frac{1}{n!} D^i(\cdot)^i_0$, which are 1 when $i = j$ and zero otherwise. But since $c$ and $\overline{c}$ are complementary, the lower left block is all zeros, while the lower right block is the identity. Hence,

$$\Lambda_{r, \overline{c}}^T V_{c, \overline{c}} = \left[ \begin{array}{c}
T_{r, c} \\
0 \\
I \end{array} \right].$$

Therefore, $\det(\Lambda_{r, \overline{c}}^T V_{c, \overline{c}}) = \det(T_{r, c})$, and so $\Lambda_{r, \overline{c}}^T V_{c, \overline{c}}$ is invertible if $T_{r, c}$ is invertible. \hfill $\square$

For example, suppose $r = [0, 2, 4, 7]$ and $c = [1, 2, 5, 8]$ in $[0 : 8]$. Then,

$$\Lambda_{r, \overline{c}} = \left[ \begin{array}{cccc}
\delta_0 & \delta_1 & \delta_2 & \delta_3 \\
\frac{1}{2} \delta_1 D^2 & \frac{1}{2} \delta_2 D^2 & \frac{1}{2} \delta_3 D^2 & \frac{1}{2} \delta_4 D^2 \\
\frac{1}{4} \delta_1 D^4 & \frac{1}{4} \delta_2 D^4 & \frac{1}{4} \delta_3 D^4 & \frac{1}{4} \delta_4 D^4 \\
\frac{1}{8} \delta_1 D^6 & \frac{1}{8} \delta_2 D^6 & \frac{1}{8} \delta_3 D^6 & \frac{1}{8} \delta_4 D^6 \\
\frac{1}{16} \delta_1 D^8 & \frac{1}{16} \delta_2 D^8 & \frac{1}{16} \delta_3 D^8 & \frac{1}{16} \delta_4 D^8 \\
\frac{1}{32} \delta_1 D^{10} & \frac{1}{32} \delta_2 D^{10} & \frac{1}{32} \delta_3 D^{10} & \frac{1}{32} \delta_4 D^{10} \end{array} \right],$$

$$V_{c, \overline{c}} = \left[ \begin{array}{ccccc}
(\cdot) & (\cdot)^2 & (\cdot)^4 & (\cdot)^8 & 1 \\
(\cdot) & (\cdot)^3 & (\cdot)^4 & (\cdot)^6 & (\cdot)^7 \end{array} \right].$$

And so,

$$\Lambda_{r, \overline{c}}^T V_{c, \overline{c}} = \begin{bmatrix}
\frac{1}{16} \delta_0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
\frac{1}{32} \delta_0 D^3 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\frac{1}{64} \delta_0 D^6 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
\frac{1}{128} \delta_0 D^7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\frac{1}{256} \delta_0 D^8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$ 

3. Birkhoff Interpolation and the Pólya condition

In the previous section, we established an equivalent condition for the invertibility of the submatrix $T_{r, \overline{c}}$ of the Pascal matrix. This equivalent condition was expressed in terms of a kind of generalized Vandermonde determined by the evaluation of certain derivatives at 0 and 1. As it turns out, this kind of interpolation problem is called (2-point) Birkhoff interpolation. To solve problems like this, Ferguson ([2]) used incidence matrices.

**Definition 3.1.** An incidence matrix $E$ for 2-point interpolation problems on $\Pi_n$ is a $2 \times (n + 1)$ matrix 

$$E := \begin{bmatrix} e_{ij} \end{bmatrix} = \begin{bmatrix} e_{00} & e_{01} & \cdots & e_{0m} \\
e_{10} & e_{11} & \cdots & e_{1m} \end{bmatrix}$$

of ones and zeros, with exactly $n + 1$ ones. The term $e_{ij}$ is 1 when the interpolation problem includes $\delta_x D^i$. 
For example, the following are incidence matrices, each of dimension $2 \times 6$ with exactly 6 ones:

$$
\begin{bmatrix}
0 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 1
\end{bmatrix},
\begin{bmatrix}
0 & 1 & 0 & 0 & 1 & 1 \\
0 & 1 & 1 & 0 & 1 & 0
\end{bmatrix}.
$$

These correspond to the functionals

$$
[\delta_{x_0}, \delta_{x_0}^3, \delta_{x_1}, \delta_{x_1}^2, \delta_{x_1}^3, \delta_{x_1}^5]
$$

and

$$
[\delta_{x_0}, \delta_{x_0}^4, \delta_{x_0}^5, \delta_{x_1}, \delta_{x_1}^2, \delta_{x_1}^4],
$$

respectively. In this paper, we take $x_0 = 0$ and $x_1 = 1$. As in [2], we define $M_j$ to be the cumulative column sum

$$
M_j := \sum_{i=0}^{1} \sum_{k=0}^{j} e_{ik}
$$

for $j = 0 : n$. In the previous examples, $M = [1, 2, 3, 5, 5, 6]$ and $M = [0, 2, 3, 3, 5, 6]$. Note that $M_5 = 6$ is the total number of ones in both examples. In general, $M_n = n + 1$ for any incidence matrix.

The problem of 2-point Birkhoff interpolation was studied by Pólya. With respect to the incidence matrices, the following definition is used.

**Definition 3.2 ([2]).** The incidence matrix $E$ satisfies the Pólya condition if $M_j > j$ for $j = 0 : m$.

In the above two examples, the first matrix is Pólya, however the second is not because $M_0 = 0 < 1$ and $M_3 = 3 < 4$. The following result, proved independently by Pólya and Whittaker (as also described in [2]), gives necessary and sufficient conditions for correct interpolation.

**Theorem 3.3** (adapted from [7], [8]). Let $E$ be a $2 \times (n + 1)$ incidence matrix with entries $e_{ij}$, ones or zeros. Let

$$
\Lambda := [\delta_{x_i} D^j : e_{ij} = 1]
$$

and $V^n = [1, (\cdot), \ldots, (\cdot)^n]$. Then, the system $\Lambda^T V^n$ is invertible iff $E$ satisfies the Pólya condition.

4. Invertibility of Submatrices of the Pascal Matrix

Suppose $r = [r_0, \ldots, r_m]$ and $c = [c_0, \ldots, c_m]$ are selections (increasing sequences) of the rows and columns of $T$, respectively. Let $\overline{c}$ be the complement of $c$ in $[0 : n]$ with $n := \max\{r_m, c_m\}$. We define $E_{r,\overline{c}}$ to be the $2 \times (n + 1)$ matrix with $E_{r,\overline{c}}(0, j) = 1$ if $r_i = 1$, and $E_{r,\overline{c}}(1, j) = 1$ if $\overline{c}_i = 1$, with all other entries 0. For example, let $r = [0, 1, 4]$ and $c = [0, 4, 5]$. We choose $n = 5$. Then, $\overline{c} = [1, 2, 3]$, and

$$
E_{r,\overline{c}} = E_{[0,1,4],[1,2,3]} = \begin{bmatrix}
1 & 1 & 0 & 0 & 1 & 0 \\
0 & 1 & 1 & 1 & 0 & 0
\end{bmatrix}.
$$

The following verifies that $E_{r,\overline{c}}$ is an incidence matrix.

**Lemma 4.1.** $E_{r,\overline{c}}$ is an incidence matrix of dimension $2 \times (n + 1)$.
Proof. Since,
\[ \#c = (n + 1) - \#\overline{c} = (n + 1) - (m + 1) = n - m, \]
it follows that
\[ \sum_{i=0}^{1} \sum_{j=0}^{n} e_{ij} = \#r + \#\overline{r} = (m + 1) + (n - m) = n + 1. \]
Hence, there are exactly \( n + 1 \) ones (and \( n + 1 \) zeros). Following Definition 3.1, \( E_{r,\overline{r}} \) is an incidence matrix. □

The next results are used to prove Theorem 1.1. The first is a Corollary of two previous theorems.

**Corollary 4.2.** The Pascal submatrix \( T_{r,c} \) is invertible iff \( E_{r,\overline{r}} \) is Pölya.

Proof. By Theorem 2.3, \( T_{r,c} \) is invertible iff \( \Lambda_{T_{r,c}} V_{c,c} \) is invertible, and by Theorem 3.3, \( \Lambda_{T_{r,c}} V_{c,c} \) is invertible iff \( E_{r,\overline{r}} \) is Pölya. Therefore, \( T_{r,c} \) is invertible iff \( E_{r,\overline{r}} \) is Pölya. □

**Lemma 4.3.** \( r \leq c \) iff \( \overline{c} \leq \overline{r} \).

Proof. Suppose \( r \leq c \). By the tautologies
\[ \overline{c} \cap c = c \cap \overline{r} \cap c, \]
we see that \( \overline{r} \) contains those elements in \( c \) that are not in both \( c \) and \( r \), and \( \overline{c} \) contains those elements in \( r \) that are not in both \( c \) and \( r \). Since \( r \leq c \), it follows that \( \overline{r} \geq \overline{c} \).

For the reverse implication, assume that \( \overline{c} \leq \overline{r} \), and apply a similar proof as above, noting that \( \overline{r} = r \) and \( \overline{c} = c \). □

**Lemma 4.4.** Let \( r \) and \( x \) be selection vectors of the same length. Then, \( E_{r,x} \) is Pölya iff \( x \leq \overline{c} \).

Proof. Suppose first that \( x = \overline{r} \). We want to establish that \( E_{r,\overline{r}} \) is Pölya. Since \( r \) and \( \overline{r} \) are complementary, each column of the incidence matrix \( E_{r,\overline{r}} \) has exactly one “1” and one “0”. Hence, \( M_0 = 1, M_1 = 1 + 1 = 2, \) etc. In particular \( M_j = j + 1 \) for \( j = 0, \ldots, n \), implying that \( E_{r,\overline{r}} \) is Pölya.

Now, suppose that \( x \leq \overline{r} \). Let \( M^x := [M_0^x, \ldots, M_n^x] \) and \( M^\overline{r} := [M_0^\overline{r}, \ldots, M_n^\overline{r}] \) be the cumulative sums defined in (1). Since \( x \leq \overline{r} \), it follows that
\[ M_j^x \geq M_j^\overline{r} = j + 1. \]
Hence, \( E_{r,x} \) is Pölya.

Now suppose that \( x \not\leq \overline{r} \). Then, \( x_i > \overline{r}_i \) for some \( i \). Suppose that \( j \) is the first such occurrence. Then, \( M_{j-1}^x = j = M_j^\overline{r} \). Hence, we don’t have \( M_j^x > j \) for all \( j \), implying that \( E_{r,x} \) is not Pölya. By contraposition, \( E_{r,x} \) Pölya implies \( x \leq \overline{r} \).

On combining these last two results, we have \( E_{r,x} \) Pölya iff \( x \leq \overline{r} \). □

We now establish our main result.

**Theorem 1.1.** Submatrices \( T_{r,c} \) of the upper Pascal matrix are invertible iff the following equivalent conditions hold:
• $r \leq c$ (i.e., $r_i \leq c_i$ for all $i$).
• There is no zero diagonal entry.

Proof. For the first part, $r \leq c$ iff $\bar{c} \leq \bar{r}$ by Lemma 4.3 and $\bar{c} \leq \bar{r}$ iff $E_{r,c}$ is Pólya by Lemma 4.4. Hence, $E_{r,c}$ is Pólya iff $r \leq c$. By Corollary 4.2, this holds iff $T_{r,c}$ is invertible. This establishes the first part.

For the second part, note that the diagonal elements of $T_{r,c}$ are of the form $\binom{c_k}{r_k}$, which are zero iff $r_k > c_k$. Hence, a diagonal element is zero iff $r \not\leq c$. By the first result, this occurs iff $T_{r,c}$ is not invertible. By contraposition, $T_{r,c}$ is invertible iff there is no zero on the diagonal.

References


Georgia Southern University, USA

E-mail address: scott.kersey@gmail.com