LAA 7456
DISK /16/7/2002

ARTICLE IN PRESS

No. of pages: 16
DTD 4.3.1/SPS

ELSEVIER

LINEAR ALGEBRA AND ITS APPLICATIONS

Linear Algebra and its Applications xx (2002) xxx–xxx
www.elsevier.com/locate/laa

A matrix approach to polynomials

Teijo Arponen

Department of Mathematics and Computer Science, University of Leicester, Leicester LE1 7RH, UK

Received 27 December 2001; accepted 30 April 2002

Submitted by R.A. Brualdi

Abstract

We present a matrix formalism to study univariate polynomials. The structure of this formalism is beautiful enough to be worth seeing on its own, yet we give (another) motivation to this by presenting three new theorems and applying the formalism to give new proofs of some known results. © 2002 Elsevier Science Inc. All rights reserved.

AMS classification: 11C08; 15A36; 11B68; 11B73

Keywords: Bernoulli polynomials; Euler polynomials; Stirling coefficients; Power sum

1. Introduction

We study univariate polynomials over a field of characteristic zero. For this we develop a matrix formalism to represent the elementary operators $Mp(x) := p(x+1)$, $\Delta p(x) := p(x+1) - p(x)$ and some others. Although these operators are certainly well known and widely used also in context of more general functions than polynomials, it seems that they have not been studied as matrices. Probably because of such approach is considered too elementary.

However, studying them as matrices reveals beautiful interplay between binomial coefficients, stirling numbers and integer vandermonde matrices. Our approach is (indeed!) elementary and requires nothing more than simple linear algebra. We present three theorems which we were unable to find from literature, despite rather extensive search through literature. Also, as further applications of our formalism, we give new proofs of some known results: Kalman's formula for power sums of consecutive integers, Euler polynomials, Verde-Star's generalized Stirlings.

E-mail address: tta1@mcs.le.ac.uk

PII: S0024-3795(02)00421-4

LAA 7456 DISK /16/7/2002 ARTICLE IN PRESS No. of pages: 16 DTD 4.3.1/SPS

Many, perhaps all, of our results could be proved in an even more elementary way by using suitably chosen induction arguments, but we believe that the structure presented here offers a new viewpoint which might turn out to be valuable also for other situations as for those described here.

We mention also the interesting book [2] which is in the same spirit as this paper, but in a sense dual direction: there are polynomials used to study linear algebra, while here we use linear algebra to study polynomials.

2. Notation and conventions

We will stick to the field $\mathbb{R}$, for convenience. Denote by $\mathbb{R}^\infty$ the direct sum of numerably many $\mathbb{R}$s, that is, the collection of $\mathbb{R}$-sequences with only finitely many nonzero elements. Clearly a polynomial $p \in \mathbb{R}[x]$ can be interpreted as an element in $\mathbb{R}^\infty$ by its coefficients, i.e. define mapping $\mathrm{ke} : \mathbb{R}[x] \to \mathbb{R}^\infty$ by

$$\mathrm{ke} : p_0 + p_1 x + p_2 x^2 + \cdots + p_N x^N \mapsto (p_0,\ p_1,\ p_2,\ \ldots,\ p_N,\ 0,\ 0,\ \ldots).$$

If $p \in \mathbb{R}[x]$, we denote by p_j the coefficient of x^j of p.

Some conventions: $0^0 := 1$. We will use the same notation, p, whether $p \in \mathbb{R}[x]$ or $p = \mathrm{ke}(p(x)) \in \mathbb{R}^\infty$. It will be clear from context which one is meant. Binomial coefficient is sometimes considered as a polynomial in the upper argument (the lower argument will always be a nonnegative integer)

$$\binom{x}{0} = 1, \quad \binom{x}{k} = \frac{1}{k!} x(x-1)(x-2)\ldots(x-k+1) \text{ for } k \in \mathbb{N}.$$

Note that

$$\binom{-1}{k} = (-1)^k$$

and more generally

$$\binom{-n}{k} = (-1)^k \binom{n+k-1}{k}.$$

The stirling numbers of the first kind, denoted here by $s(n,k)$, are defined by

$$x(x-1)\cdots(x-k+1) =: \sum_{n=0}^{k} s(n,k) x^n$$

and the stirling numbers of the second kind, denoted here by $\mathscr{S}(n,k)$, are defined by

$$x^k =: \sum_{n=0}^{k} \mathscr{S}(n,k)\, x(x-1)\cdots(x-n+1).$$

A matrix or a vector ‘is integers’ means that its elements are integers. Notation $A(:,j)$ is the jth column of A and $A(j,:)$ is the jth row of A. We will numerate the rows and columns of a matrix beginning from 0, to make the formulas simpler. That is, the top row of an $(N+1) \times (N+1)$-matrix A is $A_{00},\ A_{01},\ A_{02}, \ldots, A_{0N}$. Vectors are considered as elements of $\mathbb{R}^\infty$. That is, when two vectors of different length are added, the shorter one is filled with zeros at the end. Denote $e_n := (0, \ldots, 0, 1, 0, \ldots) \in \mathbb{R}^\infty$ with the 1 at nth place, where $n = 0, 1, \ldots$ In other words, $e_n = \mathrm{ke}(x^n) \cdot$ diag $(a_1, \ldots, a_n)$ means a diagonal matrix with elements $(a_1, \ldots, a_n)$.

3. Definitions

Fix a positive integer N. Suppose $p \in \mathbb{R}[x]$ with $\deg(p) \leqslant N$. The polynomials

$$\binom{x}{0}, \ldots, \binom{x}{N}$$

in x are linearly independent over $\mathbb{R}$, because

$$\deg \binom{x}{k} = k.$$

Hence we can write

$$p(x) =: \sum_{u=0}^{N} b_u \binom{x}{u}.$$

Definition 3.1. Mappings bin, val: $\mathbb{R}[x] \to \mathbb{R}^\infty$ by

$$\mathrm{bin}(p) := \begin{pmatrix} b_0 & b_1 & b_2 & \cdots & b_N & 0 & 0 & \ldots \end{pmatrix} \in \mathbb{R}^\infty$$
$$\mathrm{val}(p) := \begin{pmatrix} p(0) & p(1) & p(2) & \cdots & p(N) & 0 & 0 & \ldots \end{pmatrix} \in \mathbb{R}^\infty.$$

Clearly mappings ke and bin are bijective and linear. So is val when we keep N fixed and restrict val to polynomials of degree $\leqslant N$. Since we have fixed N, we can present $\mathrm{ke}(p)$, $\mathrm{bin}(p)$, $\mathrm{val}(p)$ as $N+1-$vectors for any $\deg(p) \leqslant N$.

We want a matrix M such that $\mathrm{ke}(p(x+1)) = M\,\mathrm{ke}(p(x))$. That is, M shifts the graph of a polynomial one step to the left. It turns out that (see also [3])

$$M = \begin{pmatrix} \binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \cdots & \binom{N}{0} \\ & \binom{1}{1} & \binom{2}{1} & \cdots & \binom{N}{1} \\ & & \binom{2}{2} & \cdots & \binom{N}{2} \\ & & & \ddots & \vdots \\ & & & & \binom{N}{N} \end{pmatrix}.$$

LAA 7456
DISK /16/7/2002
ARTICLE IN PRESS
No. of pages: 16
DTD 4.3.1/SPS

Indeed,

$$p(x+1)=\sum p_j(x+1)^j=\sum_j p_j \sum_{k=0}^{j}\binom{j}{k}x^k.$$

Now the coefficient of x^k is the sum over such indices j that $N \geqslant j \geqslant k$:

$$\text{coefficient of } x^k=\sum_{j=k}^{N}\binom{j}{k}p_j$$

but this is just $(Mp)_k$, i.e. the kth element of Mp.

Definition 3.2. We shall make use of the following vandermonde matrix:

$$V := \text{vandermonde}(0,\ 1,\ \ldots,\ N)$$

the stirling matrices:

$$\mathrm{St}_1 := (s(i,\ j))_{i,\ j=0\ldots N}\ ,$$
$$\mathrm{St}_2 := (\mathscr{S}(i,\ j))_{i,\ j=0\ldots N}\ .$$

Further we define, with a a scalar,

$$\Lambda := \operatorname{diag}(0,\ 1,\ \ldots,\ N),$$
$$a^{\Lambda} := \operatorname{diag}(1,\ a,\ a^2,\ \ldots,\ a^N),$$
$$\Lambda! := \operatorname{diag}(0!,\ 1!,\ 2!,\ldots,\ N!).$$

That is,

$$\mathrm{St}_1=\begin{pmatrix} s(0,0) & s(0,1) & s(0,2) & \ldots & s(0,N)\\ 0 & s(1,1) & s(1,2) & \ldots & s(1,N)\\ 0 & 0 & s(2,2) & \ldots & s(2,N)\\ \vdots & & \ddots & & \vdots\\ 0 & 0 & 0 & \ldots & s(N,N)\end{pmatrix},\ \text{Stirling 1st kind},$$

$$\mathrm{St}_2=\begin{pmatrix} \mathscr{S}(0,0) & \mathscr{S}(0,1) & \mathscr{S}(0,2) & \ldots & \mathscr{S}(0,N)\\ 0 & \mathscr{S}(1,1) & \mathscr{S}(1,2) & \ldots & \mathscr{S}(1,N)\\ 0 & 0 & \mathscr{S}(2,2) & \ldots & \mathscr{S}(2,N)\\ \vdots & & \ddots & & \vdots\\ 0 & 0 & 0 & \ldots & \mathscr{S}(N,N)\end{pmatrix},$$

Stirling 2nd kind,

$$V=\begin{pmatrix} 1 & 1 & 1 & 1 & \ldots & 1\\ 0 & 1 & 2 & 3 & \ldots & N\\ 0 & 1 & 2^2 & 3^2 & \ldots & N^2\\ \vdots & & & \ddots & & \vdots\\ 0 & 1 & 2^N & 3^N & \ldots & N^N\end{pmatrix}.$$

LAA 7456
DISK /16/7/2002

ARTICLE IN PRESS

No. of pages: 16
DTD 4.3.1/SPS

It is easily checked that $\sum_{\nu=0}^{\infty} s(n,\nu)\mathscr{S}(\nu,m) = \delta_{nm}$ (Kronecker delta) which implies $\mathrm{St}_1^{-1} = \mathrm{St}_2$.

4. Structure

We begin with a simple lemma whose proof can be omitted.

Lemma 4.1. *Suppose A is an $(N+1)-$square matrix, $p(x)$ a polynomial and $s \in \mathbb{R}$. Then*:
1. $(s^{-\Lambda} A s^{\Lambda})_{ij} = s^{j-i} A_{ij},\ s \neq 0$,
2. $(\Lambda!^{-1} A \Lambda!)_{ij} = \frac{j!}{i!} A_{ij}$,
3. $\mathrm{ke}(p(sx)) = s^{\Lambda}\,\mathrm{ke}(p(x))$,
4. $\binom{s+1}{k} = \binom{s}{k} + \binom{s}{k-1}$, *where* $k = 1, 2, \ldots$

As Kalman has noted in [3], $M = \exp(D)$ where

$$D = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 2 & \cdots & 0 \\ & & & \ddots & \vdots \\ & & & 0 & N \\ & & & & 0 \end{pmatrix}. \tag{1}$$

Note that in Kalman's definition the transpose of our M has been used. Now clearly $\mathrm{ke}(Dp) = \mathrm{ke}(p')$ where prime indicates the usual derivative. From the definition of M it follows that $M^n\,\mathrm{ke}(p(x)) = \mathrm{ke}(p(x+n))$ for all $n \in \mathbb{N}$. Moreover we have:

Proposition 4.1.

$$\mathrm{ke}(p(x+s)) = M^s \mathrm{ke}(p), \quad \textit{where } M^s := \exp(sD), \tag{2}$$

$$M^s = s^{-\Lambda} M s^{\Lambda} \quad \forall s \in \mathbb{R}\backslash\{0\} \tag{3}$$

especially, when $s \in \mathbb{Z}$, M^s coincides with the sth *power of M*.

Proof. First we need to show (3) for $s = 1$. Now D is nonzero only at its first diagonal, hence D^k is nonzero only at kth diagonal, when $k = 0, \ldots, N$ and $D^k = 0$ for $k > N$. More precisely,

$$(D^k)_{ij} = \begin{cases} j(j-1)\cdots(j-(k-1)) & \text{when } j = i+k, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

LAA 7456 DISK /16/7/2002 ARTICLE IN PRESS No. of pages: 16 DTD 4.3.1/SPS

i.e.

$$(D^k)_{ij} = j!/i!, \quad k = j - i.$$

Hence

$$(e^D)_{ij} = \left(\sum_{k=0}^{N} \frac{1}{k!} D^k\right)_{ij} = \frac{1}{(j-i)!}(D^{j-i})_{ij} = \binom{j}{i} = M_{ij}, \tag{5}$$

which shows that the claim is true for $s = 1$. By Lemma 4.1

$$s^{-\Lambda} D s^{\Lambda} = sD \quad \forall s \in \mathbb{R} \tag{6}$$

and by induction

$$s^{-\Lambda} D^k s^{\Lambda} = (s^{-\Lambda} D s^{\Lambda})(s^{-\Lambda} D^{k-1} s^{\Lambda}) = sD\, s^{k-1} D^{k-1} = s^k D^k. \tag{7}$$

Hence

$$e^{sD} = \sum_{k=0}^{N} \frac{1}{k!} s^k D^k = s^{-\Lambda} \left(\sum_{k=0}^{N} \frac{1}{k!} D^k\right) s^{\Lambda} = s^{-\Lambda} M s^{\Lambda} \tag{8}$$

which proves (3). To prove (2) note that

$$p(x+s) = \sum_{j=0}^{N} p_j \sum_{i=0}^{j} \binom{j}{i} x^i s^{j-i} = \sum_{i=0}^{N} x^i \sum_{j=i}^{N} \binom{j}{i} s^{j-i} p_j \tag{9}$$

and by Lemma 4.1,

$$\binom{j}{i} s^{j-i} = (s^{-\Lambda} M s^{\Lambda})_{ij} = (M^s)_{ij}. \tag{10}$$

Combining (9) and (10) gives $\mathrm{ke}(p(x+s))_i = \sum_{j=i}^{N} (M^s)_{ij} p_j = (M^s\, \mathrm{ke}(p(x)))_i$. The final claim follows from the fact that:

$$\exp(nA) = (\exp A)^n \quad \forall n \in \mathbb{Z}$$

for any square matrix A. $\square$

An immediate corollary is that $M^{s+t} = M^s M^t$ for all $s, t \in \mathbb{R}$. This proposition also has, by Lemma 4.1, a nice geometrical interpretation: $M^s p =$ “translate the graph of p s units” $= s^{-\Lambda} M s^{\Lambda} p =$ “zoom out s units, translate 1 unit, zoom in s units”.

Next we look at relations between ke, bin, val:

Proposition 4.2. $V^{\mathrm{T}}\, \mathrm{ke}(p) = \mathrm{val}(p) = M^{\mathrm{T}}\, \mathrm{bin}(p)$.

LAA 7456 DISK /16/7/2002 ARTICLE IN PRESS No. of pages: 16 DTD 4.3.1/SPS

Proof. Writing the definitions in suitable form:

$$(\mathrm{val}\, p)_i = p(i) = \sum_{\nu=0}^{N} (\mathrm{bin}\, p)_\nu \binom{i}{\nu} = M(:, i)\mathrm{bin}(p) = (M^{\mathrm{T}}\, \mathrm{bin}(p))_i,$$

$$(\mathrm{val}\, p)_i = p(i) = \sum_{\nu=0}^{N} (\mathrm{ke}\, p)_\nu\, i^\nu = V(:, i)\mathrm{ke}(p) = (V^{\mathrm{T}}\, \mathrm{ke}(p))_i. \qquad \square$$

Let $\varDelta$ be the difference operator: $\varDelta p(x) := p(x+1) - p(x)$. Now in ke − ke bases clearly $\varDelta = M - I$. It turns out that $\varDelta$ has a particularly simple form in bin − bin bases: it is a shift. More precisely,

Proposition 4.3. $\mathrm{bin}(\varDelta p) = S\, \mathrm{bin}(p)$ *where* S *is the shift matrix*: $S_{ij} = \delta_{i,i+1}$ (*kronecker delta*).

Proof. Clearly,

$$\varDelta \binom{x}{0} = 0$$

and by Lemma 4.1,

$$\varDelta \binom{x}{k} = \binom{x}{k-1} \qquad \forall x \in \mathbb{R}, \quad k \in \mathbb{N},$$

hence

$$\begin{aligned} \varDelta p &= \varDelta \sum_{k=0}^{N} (\mathrm{bin}\, p)_k \binom{x}{k} = \sum_{k=1}^{N} (\mathrm{bin}\, p)_k \binom{x}{k-1} \\ &= \sum_{k=0}^{N-1} (\mathrm{bin}\, p)_{k+1} \binom{x}{k}, \end{aligned} \tag{11}$$

which shows that $(\mathrm{bin}(\varDelta p))_k = (\mathrm{bin}\, p)_{k+1}$. $\square$

We have the following decompositions:

Theorem 1.
1. *The* LDU *decomposition of* $V = \mathrm{St}_2^{\mathrm{T}}\, \varLambda!\, M$,
2. *The Jordan decomposition of* $M = (\mathrm{St}_1\, \varLambda!^{-1})(I + S)(\varLambda!\, \mathrm{St}_2)$.

Proof. 1. With a fixed $k \in \{0, 1, \ldots, N\}$ define

$$p(x) := x(x-1)\ldots(x-k+1) = k! \binom{x}{k} = \sum_n s(n,k)x^n$$

as follows from definitions of $\binom{x}{k}$ and $s(n,k)$. Now $\mathrm{ke}(p) = \mathrm{St}_1(:, k)$ and

$$\mathrm{val}(p) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ k!\binom{k}{k} \\ k!\binom{k+1}{k} \\ \vdots \\ k!\binom{N}{k} \end{pmatrix} = k!\, M^{\mathrm{T}}(:,k) = (M^{\mathrm{T}}\,\Lambda!)(:,k) \tag{12}$$

on the other hand, from Proposition 4.2, $\mathrm{val}(p) = V^{\mathrm{T}}\,\mathrm{ke}(p) = (V^{\mathrm{T}}\,\mathrm{St}_1)(:,k)$, which holds for all $k \in \{0, 1, \dots, N\}$, therefore $M^{\mathrm{T}}\,\Lambda! = V^{\mathrm{T}}\,\mathrm{St}_1$. This is equivalent with $V = \mathrm{St}_2^{\mathrm{T}}\,\Lambda!\,M$ which is the LDU decomposition since both St_2 and M are upper triangulars with unit diagonal.

2. By Proposition 4.3, for p an arbitrary polynomial,

$$\begin{aligned} \mathrm{bin}\, p(x+1) &= \mathrm{bin}(p(x)) + \mathrm{bin}(\Delta p(x)) = \mathrm{bin}(p(x)) + S\,\mathrm{bin}(p(x)) \\ &= (I+S)\mathrm{bin}(p(x)) \end{aligned} \tag{13}$$

and by Proposition 4.2 and the part 1 just proven:

$$\mathrm{bin}\, p = M^{-T}\,V^{\mathrm{T}}\mathrm{ke}\, p = \Lambda!\,\mathrm{St}_2\,\mathrm{ke}\, p$$

hence from (13),

$$\Lambda!\,\mathrm{St}_2\,\mathrm{ke}\, p(x+1) = (I+S)\Lambda!\,\mathrm{St}_2\,\mathrm{ke}\, p(x) \tag{14}$$

since $M\,\mathrm{ke}(p(x)) = \mathrm{ke}(p(x+1))$ this becomes

$$\mathrm{ke}(p(x+1)) = M\,\mathrm{ke}(p(x)) = (\Lambda!\,\mathrm{St}_2)^{-1}(I+S)\Lambda!\,\mathrm{St}_2\,\mathrm{ke}\, p(x) \quad \forall p \tag{15}$$

hence M has the claimed form. □

The following proposition shows that stirling matrices intertwine difference and derivative:

Proposition 4.4. $\mathrm{St}_2\Delta = D\mathrm{St}_2$ *and* $\Delta\mathrm{St}_1 = \mathrm{St}_1 D$, *where* $\Delta = M - I$ *i.e. the matrix for the difference operator in* ke − ke *bases.*

Proof. The claims are clearly equivalent, it is enough to prove the first one.

$$\begin{aligned} \mathrm{St}_2\Delta &= \mathrm{St}_2(M-I) = \mathrm{St}_2\left(\mathrm{St}_1\,\Lambda!^{-1}(I+S-I)\Lambda!\,\mathrm{St}_2\right) \\ &= \Lambda!^{-1}\,S\,\Lambda!\,\mathrm{St}_2 = D\,\mathrm{St}_2, \end{aligned} \tag{16}$$

where the second equality follows from Theorem 1 and the last one from Lemma 4.1. □

We can use those properties to invert (up to an additive constant) Δ:

Proposition 4.5. *If q is a polynomial such that $\Delta q = p$, then, for any p given, q is uniquely defined except for an additive constant $q(0)$ and given by*

$$\mathrm{ke}(q) = \mathrm{St}_1\, \Lambda^\dagger \begin{pmatrix} q(0) \\ \mathrm{St}_2\, \mathrm{ke}(p) \end{pmatrix}, \tag{17}$$

where $\Lambda^\dagger = \mathrm{diag}(1, 1, 1/2, 1/3, \dots, 1/N)$ *i.e.* $\Lambda^\dagger\Lambda = \mathrm{diag}(0, 1, 1, \dots, 1)$. *Recall that we are interpreting vectors as elements in $\mathbb{R}^\infty$, hence the concatenated vector in right-hand side does not cause conflict with vector lengths. Note also that* $\deg(\mathrm{St}_2\, p) = \deg(p)$, *hence degree of right-hand side indeed equals* $\deg(q)$.

Proof. As above, $\Delta = A^{-1} S A$ with $A = \Lambda!\, \mathrm{St}_2$ invertible, hence $\ker \Delta = \ker S = \{\xi\, e_0 \mid \xi \in \mathbb{R}\}$ = the set of constant polynomials which proves the claimed uniqueness up to an additive constant. To prove (17), first note that

$$p = S \begin{pmatrix} \xi \\ p \end{pmatrix} \quad \forall \xi \in \mathbb{R},\ p \in \mathbb{R}^\infty.$$

Hence

$$\mathrm{ke}(\Delta\, q) = \mathrm{ke}(p) = \mathrm{St}_1\, \mathrm{St}_2\, \mathrm{ke}(p) = \mathrm{St}_1\, S \begin{pmatrix} 0 \\ \mathrm{St}_2\, \mathrm{ke}(p) \end{pmatrix}. \tag{18}$$

Noting that $S = D\, \Lambda^\dagger$ this becomes, by Proposition 4.4,

$$\mathrm{ke}(\Delta\, q) = \Delta\, \mathrm{St}_1\, \Lambda^\dagger \begin{pmatrix} 0 \\ \mathrm{St}_2\, \mathrm{ke}(p) \end{pmatrix}, \tag{19}$$

hence

$$q = \mathrm{St}_1\, \Lambda^\dagger \begin{pmatrix} 0 \\ \mathrm{St}_2\, \mathrm{ke}(p) \end{pmatrix} + \xi\, e_0 \tag{20}$$

for some scalar ξ. That $\xi = q(0)$, follows from multiplying (20) from left by $(e_0)^{\mathrm{T}}$ and noting that $q(0) = (e_0)^{\mathrm{T}}\, \mathrm{ke}\, q$ and $(e_0)^{\mathrm{T}}\, \mathrm{St}_1 = (e_0)^{\mathrm{T}}$. □

Proposition 4.6. *Suppose $k \neq 0$ a scalar. Then*

$$(M - (k+1))^{-1} = -\frac{1}{k} \mathrm{St}_1 \Lambda!^{-1} \begin{pmatrix} 1 & k^{-1} & k^{-2} & \dots & k^{-N} \\ & 1 & k^{-1} & \ddots & \\ & & 1 & \ddots & k^{-2} \\ & & & \ddots & k^{-1} \\ & & & & 1 \end{pmatrix} \Lambda!\, \mathrm{St}_2. \tag{21}$$

Proof. Applying Theorem 1 we only need to simplify $(I + S - (k+1)I)^{-1}$ which equals to $-\frac{1}{k}(I - \frac{1}{k}S)^{-1}$. Now S is nilpotent with $S^{N+1} = 0$. Hence

$$\left(I - \frac{1}{k}S\right)^{-1} = \sum_{i=0}^{N} k^{-i} S^i. \qquad \square$$

An interesting corollary is that $(M-2)^{-1}$ is integers, which is perhaps not so obvious otherwise.

Proposition 4.7. *Let p be a polynomial. Then*

$$p(\mathbb{Z}) \subset \mathbb{Z} \Leftrightarrow \operatorname{bin} p \textit{ is integers}. \tag{22}$$

Proof. If bin p is integers, then

$$p(x) = \sum_i (\operatorname{bin} p)_i \binom{x}{i} \in \mathbb{Z} \quad \forall x \in \mathbb{Z}.$$

On the other hand, from Proposition 4.2 we get $\operatorname{bin}(p) = M^{-T} \operatorname{val}(p)$ and from Proposition 4.1 $M^{-T} = (-1)^{\Lambda} M^{\mathrm{T}} (-1)^{\Lambda}$ is integers, hence $p(\mathbb{Z}) \subset \mathbb{Z} \Rightarrow \operatorname{val}(p)$ is integers $\Rightarrow \operatorname{bin}(p)$ is integers. $\square$

5. Applications

We begin with a theorem which appears to be new: this resolves a question on polynomials with a "multiplication property", by which is meant (see e.g. [1,5]) that the value of the polynomial at a point mx can be sampled from the interval $[x, x+1[$ of unit length, in the following sense:

$$p(mx) = \sum_{k=0}^{m-1} a_{k,m}\, p\left(x + \frac{k}{m}\right) \ \forall x \in \mathbb{R}, \ \ \forall m \in \mathbb{N}, \tag{23}$$

where the $a_{k,m}$ are independent of x.

Theorem 2. *Suppose given an $m \in \mathbb{N}$ and a set $\{a_0, a_1, \ldots, a_{m-1}\}$. Then there exists a polynomial p with property* (23) (*for this fixed m*) *if and only if*

$$\sum_k a_k = m^n \tag{24}$$

with some $n \in \mathbb{N}$. Moreover, $\deg(p) = n$ and if we require p to be monic, it is unique.

LAA 7456
DISK /16/7/2002

ARTICLE IN PRESS

No. of pages: 16
DTD 4.3.1/SPS

Proof. Apply ke to (23), use Lemma 4.1 and Proposition 4.1 to get

$$\begin{aligned}\mathrm{ke}\, p(mx) &= m^{\Lambda}\mathrm{ke}\, p(x) = \sum_k a_k \mathrm{ke}\, p\left(x + \frac{k}{m}\right) = \sum_k a_k M^{k/m}\mathrm{ke}\, p(x) \\ &= m^{-\Lambda}\left(\sum_k a_k M^k\right) m^{\Lambda}\mathrm{ke}(p). \end{aligned} \tag{25}$$

That is,

$$m^{-\Lambda}\left(\sum_k a_k M^k - m^{\Lambda}\right) m^{\Lambda} p = 0, \tag{26}$$

where $p := \mathrm{ke}\, p(x)$. Since $m^{\pm\Lambda}$ are invertible, the existence of p is equivalent with

$$\det\left(\sum_k a_k M^k - m^{\Lambda}\right) = 0, \tag{27}$$

which is, since M is upper triangular with unit diagonal, equivalent to

$$\det\left(\left(\sum_k a_k\right) I - m^{\Lambda}\right) = 0 \tag{28}$$

which is equivalent to $\sum_k a_k = m^n$ for some $n \in \{0, 1, \ldots, N\}$. Note that we can choose N as large as we want. This implies the first claim. To see what polynomials fulfill (26) note that $\sum_k a_k M^k - m^{\Lambda}$ is of the block matrix form

$$\begin{pmatrix} T_1 & * & * \\ & 0 & * \\ & & T_2 \end{pmatrix}, \tag{29}$$

where T_1, T_2 are upper triangulars and the zero is the nth diagonal element. Now T_2 is invertible (as is T_1) because its diagonal elements are of the form $\sum_k a_k - m^j$ with $j > n$ (respectively for T_1, with $j < n$). Denote $q := m^{\Lambda} p$ and $q_j = (\mathrm{ke}\, q)_j$. From (26) follows, by invertibility of T_2, that $q_j = 0$ for $j > n$. Since T_1 is invertible there are no more free parameters than q_n, hence we must require $q_n \neq 0$ to make q nonzero polynomial therefore $\deg(q) = n = \deg(p)$. Choosing $q_n = m^n$ (to make p monic) makes q, hence p, unique. $\square$

5.1. Euler polynomials

One way to define Euler polynomials is by the property

$$E_n(x + 1) + E_n(x) = 2x^n, \quad n = 0, 1, \ldots, \tag{30}$$

which is in matrixformalism: $(M+1)E_n = 2e_n$ and defines E_n uniquely, hence the matrix of first N Euler polynomials, E, is

$$E := \begin{pmatrix} & | & & | & & | & \\ \text{ke}\,E_0 & | & \text{ke}\,E_1 & | & \dots & | & \text{ke}\,E_N \\ & | & & | & & | & \end{pmatrix} = 2(M+1)^{-1}. \tag{31}$$

From this we can, using Proposition 4.6, deduce an easy algorithm to compute the coefficients of any E_n:

$$\text{ke}(E_n) = \text{St}_1 \Lambda!^{-1}(-2)^{\Lambda} T(-2)^{-\Lambda} \Lambda! \text{St}_2 e_n, \tag{32}$$

where T is the upper triangular with all elements equal to one.

Remark 5.1. We also could define the Euler polynomials of odd degree by using Theorem 2 with $a_{km} := (-1)^k m^n$, whence for n odd $\sum_k a_{km} = m^n$ and $E_n(x)$ is the unique monic polynomial given by Theorem 2.

5.2. *Newton's theorem*

The title refers to

Proposition 5.1.

$$\sum_k \Delta^k f(0) \binom{x}{k} = f(x) \quad \forall x \in \mathbb{R}. \tag{33}$$

This is so well known and elementary that it is probably included (as an exercise) in every textbook which introduces the difference operator, but we will prove it here just for fun, the proof is a neat example of our formalism.

Proof.

$$\text{bin} \binom{x}{k} = e_k,$$

hence claim is equivalent with $(\text{bin}\, f)_k = \Delta^k f(0)$. But, evaluating at $x = 0$ is just multiplying left by e_0^{T}, i.e. $\Delta^k f(0) = e_0^{\mathrm{T}} \text{bin}(\Delta^k f) = e_0^{\mathrm{T}} S^k \,\text{bin}\, f = (S^k \,\text{bin}\, f)_0 = (\text{bin}\, f)_k$. $\square$

5.3. *Bernoulli polynomials*

There is a vast amount of literature for Bernoulli polynomials. We shall prove, as an example to use our formalism, a property which appears to be new despite its elementarity. Among several equivalent definitions we choose the following: the *n*th Bernoulli polynomial, B_n, is the unique polynomial satisfying

$$B_n(x+1) - B_n(x) = nx^{n-1}, \tag{34}$$

$$B_n'(x) = nB_{n-1}(x). \tag{35}$$

That B_n is unique will be shown shortly. Moreover, we define

$$B := \left(\begin{array}{c|c|c|c} \text{ke } B_0 & \text{ke } B_1 & \dots & \text{ke } B_N \end{array} \right). \tag{36}$$

Theorem 3. *The Bernoulli matrix B defined above is the unique matrix which maps differences to derivatives and commutes with both. That is,*

$$B\varDelta = D,$$
$$DB = BD,$$
$$\varDelta B = B\varDelta,$$

where $\varDelta = M - I$, as in Proposition 4.4.

Proof. We put the definitions in matrix formalism using ke. First, $nx^{n-1} = \mathrm{d}/\mathrm{d}x\, x^n$, hence (34) says that $\varDelta B_n = De_n$ and therefore $\varDelta B = D$. As noted in the proof of Proposition 4.5, this defines B_n except for its 0. order term. Then, the left-hand side of (35) is DB_n, so the left-hand side is DB when we look at all columns of B (that is, all n) simultaneosly. How about the right-hand side? Again, looking at all n simultaneosly, its *n*th column is nB_{n-1}, which means that, firstly, the columns of B are shifted one step to right, that is multiplying by S on the right (since $(XS)(:, n) = X(:, n-1)$ for any matrix X), and, secondly, *n*th column is multiplied by n, which is multiplying by $\varLambda$ from the right. Hence (35) says $DB = BS\varLambda$. Noting that $S\varLambda = D$ gives $DB = BD$. Now evaluating (35) at $x = 0$, which is in matrixformalism just multiplying left by e_0^{T}, we get $B_{n-1}(0) = \frac{1}{n} B_{n1}$ for $n \geqslant 1$. Herefrom we see that this defines the top row of B, hence B is unique. Especially, since $\varDelta$ and D are upper triangular and the 1^{st} diagonal of $\varDelta$ is equal to the 1^{st} diagonal of D, we see that B is upper triangular with unit diagonal. We still need to show that B commutes with $\varDelta$. Now $\varDelta BB = DB = BD = B\varDelta B$ and B is invertible, hence $\varDelta B = B\varDelta$ and $\varDelta B = D$ becomes $B\varDelta = D$. □

Remark 5.2. B_n could also be defined by using Theorem 2 with $a_{k,m} := m^{n-1}$. This is actually the starting point used in [5]. In context of Bernoulli polynomials, Theorem 2 has been generalized, see [1].

5.4. Kalman's formula for the sum of consecutive powers

There is, like in the Bernoulli case, a vast amount of literature on evaluating the sum

$$p(n) := \sum_{k=1}^{n} k^r$$

with $n \in \mathbb{N}$. Quite a common nickname for this is "*the* power sum". To author's knowledge, studying this was the original motivation why Bernoulli discovered his polynomials. We are interested in the formula that Kalman proved in [4]. We state it here in our notation (Kalman also uses the transpose of this):

Proposition 5.2. *Suppose p is a polynomial with $\deg(p) = r + 1$. Then*

$$p(x) - p(0) = \begin{pmatrix} \Delta p(0) & \Delta p(1) & \dots & \Delta p(r) \end{pmatrix} M(0 : r, 0 : r)^{-1} \begin{pmatrix} \binom{x}{1} \\ \vdots \\ \binom{x}{r+1} \end{pmatrix}. \tag{37}$$

Now the power sum is a special case (as Kalman noted), take $p(n) := \sum_{k=1}^{n} k^r$. Then $p(0) = 0$, $(\Delta p)(n) = (n + 1)^r$ and deg$(p) = r + 1$, and the identity above (take a transpose of it) becomes

$$p(n) = \sum_{k=1}^{n} k^r = \left[\binom{n}{1} \dots \binom{n}{r+1}\right] M(0 : r, 0 : r)^{-T} \begin{pmatrix} 1^r \\ 2^r \\ \vdots \\ (r+1)^r \end{pmatrix}.$$

Proof. Let $N = r + 1$. Evaluating p at x:

$$p(x) = \mathrm{bin}(p)^{\mathrm{T}} \begin{pmatrix} \binom{x}{0} \\ \binom{x}{1} \\ \vdots \\ \binom{x}{N} \end{pmatrix} = p(0) + \mathrm{bin}(p)^{\mathrm{T}}_{1:N} \begin{pmatrix} \binom{x}{1} \\ \vdots \\ \binom{x}{N} \end{pmatrix}.$$

Now, by Propositions 4.3 and 4.2, $\operatorname{bin}(p)_{1:N} = S\operatorname{bin}(p) = \operatorname{bin}(\Delta p) = M^{-T}\operatorname{val}(\Delta p)$ hence

$$p(x) - p(0) = \operatorname{bin}(\Delta p)^{\mathrm{T}} \begin{pmatrix} \binom{x}{1} \\ \vdots \\ \binom{x}{N} \end{pmatrix} = \operatorname{val}(\Delta p)^{\mathrm{T}} M^{-1} \begin{pmatrix} \binom{x}{1} \\ \vdots \\ \binom{x}{N} \end{pmatrix}. \tag{38}$$

We remind the reader of our $\mathbb{R}^\infty$-convention: since M^{-1} is of size $(N+1)\times(N+1)$, the multiplication

$$M^{-1}\begin{pmatrix} \binom{x}{1} & \dots & \binom{x}{N} \end{pmatrix}^{\mathrm{T}}$$

is by definition equal to

$$M^{-1}\begin{pmatrix} \binom{x}{1} & \dots & \binom{x}{N} & 0 \end{pmatrix}^{\mathrm{T}}.$$

Now the last element of

$$M^{-1}\begin{pmatrix} \binom{x}{1} & \dots & \binom{x}{N} & 0 \end{pmatrix}^{\mathrm{T}}$$

is zero since M^{-1} is upper triangular. Therefore the last element of $\operatorname{val}\Delta p$, as well as the last row and last column of M^{-1} do not affect the value of $p(x) - p(0)$ in (38) and can be replaced by zero, hence we get (37). □

5.5. *Verde-Star's generalized Stirlings*

In [6] Verde-Star defines generalized binomial coefficients (denote g.b.c): they are combinatorial objects associated to some given sequence. For example, the usual binomial coefficients are g.b.c's associated to the sequence $1, 1, 1, \dots$ and the Stirling numbers of 2nd kind are g.b.c's associated to $0, 1, 2, \dots$ Further he defines "generalized Stirling numbers of the second kind" as the g.b.c's associated to the sequence $\{a + bk\}_{k=0}^{\infty}$ and shows [6, (6.21)] that they are equivalently defined as:

$$\mathscr{S}_{a,b}(n, m) = \sum_{k=0}^{m} \binom{m}{k} b^{m-k} a^{k-n} \mathscr{S}(n, k) \tag{39}$$

and he proves [6, (6.26)] a generalization of this:

$$\mathscr{S}_{ac,ad+b}(n, m) = \sum_{k=0}^{m} \binom{m}{k} b^{m-k} a^{k-n} \mathscr{S}_{c,d}(n, k). \tag{40}$$

We shall give a proof of this by writing them in our formalism: first, by Lemma 4.1,

$$\binom{m}{k} b^{m-k} = (M^b)_{km}$$

and $a^{k-n}\mathscr{S}(n,k) = (a^{-\Lambda}\,\mathrm{St}_2\,a^{\Lambda})_{nk}$ hence (39) says,

$$\mathscr{S}_{a,b}(n,m) = \sum_{k=0}^{m}(M^b)_{km}(a^{-\Lambda}\mathrm{St}_2a^{\Lambda})_{nk} = (a^{-\Lambda}\mathrm{St}_2a^{\Lambda}M^b)_{nm} \tag{41}$$

hence

$$S_{a,b} := \text{the matrix of } \mathscr{S}_{a,b}(n,m)\text{'s} = a^{-\Lambda}\mathrm{St}_2a^{\Lambda}M^b. \tag{42}$$

From this it is immediate to conclude (40), since $M^{ad+b} = M^{ad}M^b$:

$$\begin{aligned} S_{ac,ad+b} &= (ac)^{-\Lambda}\mathrm{St}_2(ac)^{\Lambda}M^{ad+b} \\ &= (a^{-\Lambda}c^{-\Lambda})\mathrm{St}_2(c^{\Lambda}a^{\Lambda})(a^{-\Lambda}d^{-\Lambda}Md^{\Lambda}a^{\Lambda})M^b \\ &= a^{-\Lambda}S_{c,d}a^{\Lambda}M^b, \end{aligned} \tag{43}$$

which is exactly (40), by the same reasoning as above: now

$$(a^{-\Lambda}S_{c,d}a^{\Lambda})_{nk} = a^{k-n}\mathscr{S}_{c,d}(n,k).$$

Further, $S_{1,b}$ is called “shifted Stirling numbers of the second kind” which is a convenient name, since $S_{1,b} = \mathrm{St}_2M^b$ and M^b indeed is a shift, in the sense that M^b ke $p(x) =$ ke $p(x+b)$, i.e. it translates the graph of a polynomial b steps to the left (even when $b \notin \mathbb{Z}$).

We emphasize that these results are by no means any ‘main theorem’ of [6], but we have included here a new proof of them as an example of using our formalism.

Acknowledgements

This research was supported by the Academy of Finland, project number 53080. The author is also grateful for the referee for useful comments.

References

[1] L. Carlitz, A characterization of the Bernoulli and Euler polynomials, Rend. Sem. Mat. Univ. Padova 62 (1980) 309–318.
[2] P. Fuhrmann, A Polynomial Approach to Linear Algebra, Springer, New York, 1996.
[3] D. Kalman, Polynomial translation groups, Math. Mag. 56 (1) (1983) 23–25.
[4] D. Kalman, Sums of powers by matrix methods, Fibonacci Quart. 28 (1) (1999) 60–71.
[5] D. Lehmer, A new approach to Bernoulli polynomials, Amer. Math. Mon. 95 (10) (1988) 905–911.
[6] L. Verde-Star, Interpolation and combinatorial functions, Stud. Appl. Math. 79 (1988) 65–92.