The Determinant of a Hessenberg Matrix and Some Applications in Discrete Mathematics

Ulrich Tamm

Department of Mathematics, University of Bielefeld, P. O. Box 100131
33501 Bielefeld, Germany, email tamm@mathematik.uni-bielefeld.de

It is well known that many mathematical identities can be expressed as special determinants. In this short note we shall present an explicit (not recursive) formula for the determinant of a Hessenberg matrix (in normalized form), which we could not find in literature, and point out by several examples its use in various fields of mathematics, as Geometry, Number Theory, Probability Theory, and Combinatorics. Especially, it is related to generating functions of the form $H(z) = \frac{1}{1-G(z)}$.

A *Hessenberg matrix* $A_n = (a_{i,j})_i, j = 1, \dots n$ is a quadratic matrix of the form

$$A_n = \begin{pmatrix} a_0^{(1)} & 1 & 0 & \cdots & & 0 \\ a_0^{(2)} & a_1^{(1)} & 1 & 0 & \cdots & 0 \\ a_0^{(3)} & a_1^{(2)} & a_2^{(1)} & 1 & \cdots & 0 \\ \vdots & & & & & \vdots \\ a_0^{(n-1)} & a_1^{(n-2)} & a_2^{(n-3)} & \cdots & a_{n-2}^{(1)} & 1 \\ a_0^{(n)} & a_1^{(n-1)} & a_2^{(n-2)} & \cdots & a_{n-2}^{(2)} & a_{n-1}^{(1)} \end{pmatrix} \qquad (1)$$

So $a_{i,j} = 0$ for $j > i+1$. We present the concept of the Hessenberg matrix in a normalized form , i.e. $a_{i,i+1} = 1$ for $i = 1, \dots, n$. Usually arbitrary coefficients $a_{i,i+1}$ on the upper side diagonal are allowed. Further, by A'_n is denoted the matrix obtained from A_n by replacing the 1's on the upper side diagonal with -1's, i.e. $a_{i,i+1} = -1$ for $i = 1, \dots, n$

Hessenberg matrices play an important role in numerical mathematics, because their determinants can be calculated very fast by the following recursion (for the normalized case with $d_n = \det A_n$)

$$\det A_n = d_n = \sum_{t=1}^{n} (-1)^{t-1} a_{n-t}^{(t)} d_{t-1} \qquad (2)$$

Since there also exist fast algorithms, which allow to transform a given quadratic matrix to a Hessenberg matrix, the evaluation of the characteristic polynomial of an arbitrary quadratic matrix can be reduced to the same problem for an equivalent Hessenberg matrix (cf. [6], pp. 251 – 258, where also the recursion (2) is derived). Special determinants of this form are discussed in [11] under the name recurrents and in the theory of continued fractions as continuants.

By analyzing the above recursion (2) or by evaluating the determinant by the last row, the following proposition is immediate.

Proposition. The determinant of a Hessenberg matrix A_n defined as above is

$$\det A_n = \sum_{k=1}^{n}(-1)^k \sum_{t_1+\ldots+t_k=n} a_0^{(t_1)} a_{t_1}^{(t_2)} a_{t_1+t_2}^{(t_3)} \cdots a_{t_1+\ldots+t_{n-1}}^{(t_n)} \tag{3}$$

$$\det A'_n = \sum_{k=1}^{n} \sum_{t_1+\ldots+t_k=n} a_0^{(t_1)} a_{t_1}^{(t_2)} a_{t_1+t_2}^{(t_3)} \cdots a_{t_1+\ldots+t_{n-1}}^{(t_n)} \tag{4}$$

Remark. It can also easily be verified that $\det(A_n)$ equals the permanent of the matrix A'_n.

This determinant is quite interesting, since many formulae, which occur in various fields of mathematics, have this special form.

Example 1.

$$\int_0^{a_0}\int_{x_1}^{a_1}\int_{x_2}^{a_2}\cdots\int_{x_{n-1}}^{a_{n-1}} dx_n dx_{n-1}\cdots dx_2 dx_1 = \sum_{k=1}^{n}(-1)^k \sum_{t_1+\ldots+t_k=n} \frac{a_0^{t_1}}{t_1!}\frac{a_{t_1}^{t_2}}{t_2!}\frac{a_{t_1+t_2}^{t_3}}{t_3!}\cdots\frac{a_{t_1+\ldots+t_{n-1}}^{t_n}}{t_n!}$$

which is the determinant of a Hessenberg matrix A_n as in (1) with $a_i^{(j)} = \frac{a_i^j}{j!}$ Observe that for the special choice $a_i = 1$ for all i this is just the volume of an n–dimensional simplex. For arbitrary choice of the $a_i > 0$ the volume of a more general n–dimensional body is obtained.

For the special choice $a_i^{(j)} = b_j$ for all i, i.e. on each diagonal all the elements are the same, we have

$$\det A_n = \begin{pmatrix} b_1 & 1 & 0 & \cdots & & 0 \\ b_2 & b_1 & 1 & 0 & \cdots & 0 \\ b_3 & b_2 & b_1 & 1 & \cdots & 0 \\ \vdots & & & & & \vdots \\ b_{n-1} & b_{n-2} & b_{n-3} & \cdots & b_1 & 1 \\ b_n & b_{n-1} & b_{n-2} & \cdots & b_2 & b_1 \end{pmatrix} = \sum_{k=1}^{n}(-1)^k \sum_{t_1+\ldots+t_k=n} b_{t_1}\cdots b_{t_n} \tag{5}$$

Example 2. As already pointed out above, if we set $a_i = 1$ for all i in Example 1, we obtain the volume $\frac{1}{n!}$ of an n–dimensional simplex, which can be represented as such a determinant with $b_j = \frac{1}{j!}$

Example 3. By choosing $b_j = \frac{1}{(j+1)!}$ in (5) we obtain the following nice identity for the Bernoulli numbers B_n recursively defined by $B_0 = 1, B_{n+1} = -\frac{1}{n+2}\cdot\sum_{i=0}^{n}\binom{n+2}{i}\cdot B_i$

$$B_n = (-1)^{n-1}(2n)!\cdot\det A_n$$

This example is taken from the book [11], where in Chapter 21 many more examples for determinants of Hessenberg matrices of the form (5) (under the name recurrents) are presented.

Example 4 ([7]).

$$\sum_{x_1=0}^{p_1} \sum_{x_2=x_1}^{p_2} \cdots \sum_{x_n=x_{n-1}}^{p_n} 1 = \det(\binom{p_i+1}{j-i+1})_{i,j=1,\ldots,n}$$

is the number of sequences of nonnegative integers $(q_1, \ldots, q_n)$ majorized by the sequence $(p_1, \ldots, p_n)$. Majorization or (domination) means that $\sum_{i=1}^n q_i = \sum_{i=1}^n p_i$ and $\sum_{i=1}^t q_i \leq \sum_{i=1}^t p_i$ for all $t = 1, \ldots, n-1$. For an intensive treatment of this topic see [9]. Another application of Hessenberg matrices in the theory of majorization can be found in [2].

The determinant of a Hessenberg matrix of the form (5) also arises in Combinatorial Theory (enumeration of trees [13] and lattice paths [10], theory of partitions) and Probability Theory (renewal theory, random walks) because of its connection to generating functions of the form $H(z) = \frac{1}{1-G(z)}$ (or $H(z) = \frac{G(z)}{1-G(z)}$ which we shall use later on).

Let there be given two sequences $(h_n)_{n=1,2,\ldots}$ and $(g_n)_{n=1,2,\ldots}$, where the elements b_n are obtained from the first sequence by

$$h_n = \sum_{k=1}^{n} \sum_{t_1+\ldots+t_k=n} g_{t_1} \cdots g_{t_n}$$

It is easy to verify the inversion

$$g_n = \sum_{k=1}^{n} (-1)^k \sum_{t_1+\ldots+t_k=n} h_{t_1} \cdots h_{t_n}$$

which is of the form (2). Now if

$$G(z) = g_1 z + g_2 z^2 + g_3 z^3 + \ldots$$

is the generating function for the g_i's then the generating function $H(z) = h_1 z + h_2 z^2 + h_3 z^3 + \ldots$ for the coefficients $h_i, i = 1, 2, 3, \ldots$ is obtained from $G(z)$ via

$$H(z) = \sum_{k=1}^{\infty} G(z)^k = \frac{G(z)}{1-G(z)}$$

It is clear that $g_n = \det A_n$, when A_n is defined in the form (5) (with h_i taking the role of b_i).

Such generating functions play an important role in the analysis of random walks, when $G(z)$ is the generating functions for the steps of a random walk and $H(z)$ is the generating function for the paths arising from this set of steps (cf. [8]). In [5], p. 311, $H(z)$, the probability generating function for occurence of an event is related to $G(z)$, the probability generating function for the first occurence of an event, as above.

Example 5. The standard example in this context are the Catalan numbers $C_j = \frac{1}{j+1}\binom{2j}{j}, j = 0, 1, \ldots$.
For instance, if $g_j = C_j$ is the j–th Catalan number, then $h_i = \binom{2i}{i}$.
It can also be shown that if $g_j = C_{j-1}$ is chosen as the $(j-1)$–st Catalan number (hence $g_1 = C_0 = 1$ and $g_2 = C_1 = 1$), then $h_i = C_i$ is the i-th Catalan number, which follows immediately from the well – known functional equation $G(z) = 1 + zG(z)$ for the generating function of the Catalan numbers.
Example 6. For $g_j = \binom{j+l-1}{l}$ where $l \in \{1, 2, 3, \ldots\}$ it is $h_i = \sum_t (\binom{(l+1)i-lt}{t})$.
The last formula arises as the solution of the so–called "Simon Newcomb's Problem" in the theory of partitions (e. g. [1]). Observe that for $l = 1$, i. e. $a_j = j$, the even Fibonacci numbers $b_1 = 1, b_2 = 3, b_3 = 8, \ldots$ occur.

Example 7. A special Hessenberg matrix is a tridiagonal matrix $A_n = (a_{ij})_{i,j=1,\ldots n}$ with $a_{i,j} = 0$ for $|i-j| > 1$, whose determinant d_n is usually evaluated by exploiting the three – term recurrence

$$d_n = a_{n,n} \cdot d_{n-1} - a_{n,n-1} \cdot a_{n-1,n} \cdot d_{n-2}.$$

A three – term recurrence occurs, for instance, in the theory of continued fractions and of orthogonal polynomials, which can hence be expressed as determinants, e. g. the continuant corresponding to a continued fraction (cf. [12], p. 11) is a determinant of a matrix A'_n (with -1's on the upper side diagonal). A very useful property of a sequence $(P_n)_n$ of orthogonal polynomials is that their eigenvalues interlace. This extends to Hermitian Hessenberg matrices. For a recent result on interlacing properties in the context of Hessenberg matrices see [3].
The explicit form of the determinant (2), e. g., allows to calculate directly (without using the above recurrence) the following identities for the Chebyshev polynomials $(t_n(x))_{n=1,2,\ldots}$ and the Chebychev polynomials of the second kind $(u_n(x))_{n=1,2,\ldots}$ (cf. e. g. [4], p. 228).

$$t_n(x) = \frac{n}{2} \cdot \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^i}{n-i} \binom{n-i}{i} (2x)^{n-2i} = \det T_n(x)$$

with

$$T_n(x) = \begin{pmatrix} x & 1 & 0 & \ldots & 0 & 0 & 0 \\ 1 & 2x & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 2x & 1 \\ 0 & 0 & 0 & \ldots & 0 & 1 & 2x \end{pmatrix},$$

$$u_n(x) = \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^i \binom{n-i}{i} (2x)^{n-2i} = \det U_n(x)$$

with

$$U_n(x) = \begin{pmatrix} 2x & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 2x & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 2x & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 2x \end{pmatrix}.$$

References

[1] G. E. Andrews, The Theory of Partitions, Encyclopaedia of Mathematics and its Applications, Vol. 2, Addison Wesley (1976).

[2] R. A. Brualdi and S. G. Hwang, Vector majorization via Hessenberg matrices, J. London Math. Soc. (2) 53 (1996), no. 1, 28–38.

[3] A. Bunse-Gerstner and C. Y. He, On a Sturm sequence of polynomials for unitary Hessenberg matrices, SIAM J. Matrix Anal. Appl. 16 (1995), no. 4, 1043–1055.

[4] H. S. M. Coxeter, Regular Polytopes, Dover (1973).

[5] W. Feller, An Introduction to Probability Theory and its Applications I, Wiley (1950).

[6] J. N. Franklin, Matrix Theory, Prentice Hall (1968).

[7] B. R. Handa and S. G. Mohanty, On a property of lattice paths, J. Statist. Plann. Inference 14 (1986), 59 – 62.

[8] I. Gessel, A factorization of formal Laurent series and lattice path enumeration, J. Combin. Theory A 28 (1980), 321 – 337.

[9] A. W. Marshall and I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic Press (1979).

[10] S. G. Mohanty, Lattice Path Counting and Applications, Academic Press (1979).

[11] T. Muir, Theory of Determinants, Dover (1960).

[12] O. Perron, Die Lehre von den Kettenbrüchen, Teubner (1929).

[13] R. Sedgewick and P. Flajolet, An Introduction to the Analysis of Algorithms, Addison Wesley (1996).