On computing the factors of generalized Vandermonde determinants

STEFANO DE MARCHI
Department of Mathematics and Computer Science,
University of Udine
Via delle Scienze, 206 - 33100 Udine
ITALY
demarchi@dimi.uniud.it, http://www.dimi.uniud.it/~demarchi

Abtsract: - For generalized Vandermonde determinants of the form:

$$V_{s;\ell(\mu)}(x_1, ..., x_s) = |x_i^{\mu_k}|\,, \quad i,k = 1, ..., s\ ,$$

we present two algorithms for determining the coefficients of the polynomials in which they can be factored. We assume that the x_i are distinct points in $[a,b]$, the index s stands for the polynomial degree, the multiindex μ consists of ordered integers $0 \le \mu_1 < \mu_2 < \cdots < \mu_s$ and $\ell(\mu)$ is its length μ, that is the number of non-zero elements.

These determinants can be expressed as a product of the classical Vandermonde determinant and a homogeneous symmetric function, i.e. a *Schur function*. Moreover, a n-degree Schur function can be factored as $(x-x_1)\cdots(x-x_k)\cdot P_{n-k}(x)$. Algorithms for the computation of the coefficients of P_{n-k} are studied.

Key-Words: - Interpolation, Generalized Vandermonde determinants, Schur functions.

1 Introduction

Let $X_s = \{x_1, ..., x_s\}$ be a set of $s \ge 2$ *distinct points* in $[a,b]$. By VDM_s we indicate the *classical* Vandermonde determinant of order s, that is:

$$VDM_s = \left| x_i^{j-1} \right|_{i,j=1,\ldots,s} .$$

It is well-known that (cf. [1]):

$$VDM_s = VDM_{s-1}\left(\prod_{1\le i\le s-1}(x_s - x_i)\right) . \quad (1)$$

Let $V_{s-1}(x)$ be the polynomial of degree $s-1$ obtained when $x_s = x$. Then $V_{s-1}(x)$ (cf. (1)) can be factored as the product of the classical Vandermonde of order $s-1$ times a polynomial of degree $s-1$.

By this idea we may compute *generalized* Vandermonde determinants of the form:

$$V_{n;\ell(\mu)}(x_1, ..., x_n) = |x_i^{\mu_k}|\,, \quad i,k = 1, ..., n\ , \quad (2)$$

where the x_i are n distinct points belonging to an interval $[a,b]$ of the real line, the index n stands for the degree, the multiindex μ consists of ordered integers $0 \le \mu_1 < \mu_2 < \cdots < \mu_n$ and $\ell(\mu)$ is its length μ, that is the number of non-zero elements. This multiindex defines the functions basis of the determinant. That is, the classical Vandermonde determinant of order n is based on the polynomial basis $\{1, x, ..., x^n\}$ while the generalized ones are based on the basis $\{x^{\mu_1}, ..., x^{\mu_n}\}$ for some positive integers $0 \leq \mu_1 < \mu_2 < \cdots < \mu_n$.

Generalized Vandermonde determinants are connected to *Schur functions* (cf. [4, 5]). In fact, they can be factored as a product of the classical Vandermonde determinant and a *Schur function* (see Section 3, Theorem 3.1):

$$V_{n;\ell(\mu)}(x_1, ..., x_n) = VDM_n \cdot s_{m_{n-1}}(x_1, ..., x_n) ,$$

where $m_{n-1} = \mu_n - \mu_1$. We prove that

$$s_{m_{n-1}}(x) = \prod_{i=1}^{n-1} (x - x_i) \cdot P_M(x) , \tag{3}$$

with $M = m_{n-1} - n + 1$ when x is taken instead of x_n. Our main concern is then the computation of the coefficients of the polynomial $P_M(x)$.

2 A first class of generalized Vandermonde determinants

In the end of the twenties, Heineman [3], proved some interesting formulas for different types of generalized Vandermonde determinants, but he did not relate them to Schur functions, even if he discovered that every generalized Vandermonde determinant is an integral rational function of certain special determinants.

Let us start our investigations considering the function (determinant)

$$V_{n;k}(x) = \begin{vmatrix} 1 & x_1 & \dots & x_1^{k-1} & x_1^n \\ 1 & x_2 & \dots & x_2^{k-1} & x_2^n \\ \vdots & & \vdots & & \vdots \\ 1 & x_k & \dots & x_k^{k-1} & x_k^n \\ 1 & x & \dots & x^{k-1} & x^n \end{vmatrix} , \tag{4}$$

when the $n - k$ powers $k, ..., n-1$ are missing, $2 < k < n - 1$.

Obviously

$$V_{n;k}(x) = VDM_k \ (x - x_1) \cdots (x - x_k) \ P_{n-k}(x) . \tag{5}$$

In order to compute the coefficients of the polynomial $P_{n-k}(x)$ we divide in (5) by the constant factor VDM_k. We observe that the coefficients of the divisor $Q_k(x) = \prod_{i=1}^{k}(x - x_i)$ are, up to the sign, the `elementary symmetric functions` of the points x_i, $i = 1, ..., k$, which are usually defined as follows.

$$\begin{aligned} e_0^{(k)} &= 1 \\ e_r^{(k)} &= \sum_{i_1 < \ldots < i_k} x_{i_1} \cdots x_{i_k}, \quad 1 \leq r \leq k \\ e_r^{(k)} &= 0, \quad r < 0, \quad \text{or} \quad r > k . \end{aligned}$$

This latter is normally assumed *by definition*.

In [2], we proved that this process is equivalent to the solution of a linear system whose matrix is a *lower triangular Toeplitz* matrix generated by the vector of the elementary symmetric functions. The solution of the problem is simply obtained by finding the inverse of this matrix. If a lower triangular Toeplitz is invertible, then the inverse is still a lower triangular Toeplitz matrix. Hence, the inverse of an invertible lower triangular Toeplitz matrix is completely defined by the knowledge for example of its first column. Given the lower triangular $n \times n$ Toeplitz matrix generated by the vector $\mathbf{a} = (1, a_1, a_2, ..., a_{n-1})^T$, the *first* column of the inverse A^{-1}, that is the vector $\mathbf{b} = (b_1, ..., b_n)^T$, is found by solving the system

$$A\mathbf{b} = \mathbf{e}_1$$

where $\mathbf{e}_1 = (1, 0, ..., 0)^T$. The solution provides the following useful recurrence:

$$\begin{aligned} b_{1,1} &= 1 ; \\ b_{i,1} &= -\sum_{j=1}^{i-1} b_{j,1} a_{i-j}, \quad i = 2, ..., n . \end{aligned} \tag{6}$$

Hence, an algorithm that allows us to compute the coefficients of the quotient polynomial $P_{n-k}(x)$ can be formulated as follows.

Algorithm 1.

`Inputs:` $n,\ k$ and the points $x_i,\ i = 1, ..., k$.

1.Form the $n-k+1$ dimensional array $\mathbf{a}$, that is $\mathbf{a} = (\underbrace{e_0^{(k)}}_{=1}, -e_1^{(k)}, e_2^{(k)}, ..., (-1)^k\ e_k^{(k)}, 0, ..., 0)$;
where the latter $n-2k$ zeros correspond to the symmetric functions $e_{k+1}^{(k)}, \dots, e_{n-k}^{(k)}$.
2. By means of the recurrence (6) compute the array $\mathbf{b}$.

`Output:` the array $\mathbf{b}$ of the coefficients of $P_{n-k}(x)$.

For an analysis of the complexity of this algorithm, we refer the reader to [2, §3].

Using the `complete elementary symmetric functions` we can go further. Given q distinct points, the homogeneous symmetric polynomials

$$\begin{aligned} h_s^{(q)} &= \sum_{|\lambda|=s} m_\lambda \\ h_s^{(q)} &= 0, \quad s<0, \quad \text{or} \quad s>q\ . \end{aligned}$$

are the complete elementary symmetric functions. Obviuolsy, $h_0^{(q)} = e_0^{(q)}$, $h_1^{(q)} = e_1^{(q)}$ and $h_s^{(q)}$ is the sum of all monomials of total degree s.

Finally, the following result holds.

Theorem 2.1 *Given k distinct points, the coefficients of the polynomial $P_{n-k}(x)$ in (5), i.e. the elements of the array* $\mathbf{b}$*, are the complete elementary symmetric functions:*

$$\mathbf{b}_i = h_{i-1}^{(k)}\ , \quad i = 1, ..., n-k+1.$$

Proof. Let us consider the array $\mathbf{a}$ as above, made up of the elementary symmetric functions of the points $x_1, ..., x_k$. From the recurrence (6), we have

$$\mathbf{b}_i = \det \begin{pmatrix} e_1^{(k)} & e_2^{(k)} & \dots & \dots & e_{i-1}^{(k)} \\ e_0^{(k)} & e_1^{(k)} & \dots & \dots & e_{i-2}^{(k)} \\ e_{-1}^{(k)} & e_0^{(k)} & e_1^{(k)} & \dots & e_{i-3}^{(k)} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ e_{-i-2}^{(k)} & \dots & \dots & e_0^{(k)} & e_1^{(k)} \end{pmatrix},$$

which is $h_{i-1}^{(k)}$. □

Corollary 2.2 *For the polynomial $P_{n-k}(x)$ holds the determinantal form:*

$$P_{n-k}(x) = \\ = \begin{vmatrix} e_0^{(k)} & -e_1^{(k)} & \dots & (-1)^{n-k} e_{n-k}^{(k)} \\ 0 & e_0^{(k)} & \dots & (-1)^{n-k-1} e_{n-k-1}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -e_1^{(k)} \\ 1 & x & \dots & x^{n-k} \end{vmatrix}.$$

3 Schur functions and the general setting

We start by introducing the definition of *Schur functions*, while for a complete treatise of Schur functions we refer the reader to (cf. [5]).

Definition 3.1 *Given a partition $\lambda = (\lambda_1, ..., \lambda_n)$ of length $|\lambda| \le n$, the Schur function s_λ, defined on $I\!R^n$, is the following ratio*

$$s_\lambda(x_1, ..., x_n) = \frac{det\ (x_i^{\lambda_j+n-j})}{det\ (x_i^{n-j})}, \quad 1 \le i, j \le n. \tag{7}$$

We note that the denominator is the classical Vandermonde determinant VDM_n.

Therefore, to fit our setting in this framework, we simply have to substitute the last variable involved x_n, with x.

General Vandermonde determinants can be described as follows. Given a multiindex, $0 \leq \mu_1 \leq \mu_2 \leq \ldots \leq \mu_n$, and a set of n distinct points, $x_1, ..., x_n$, we consider determinants of the form

$$V_{\mu_n;\ell(\mu)} = \left| x_i^{\mu_j} \right|_{1 \leq i,j \leq n} . \qquad (8)$$

These determinants can be rewritten

$$V_{\mu_n;\ell(\mu)} = \left(\prod_{i=1}^{n} x_i^{\mu_1} \right) \det \left(x_i^{\mu_j - \mu_1} \right)_{1 \leq i,j \leq n} , \qquad (9)$$

and so for our purposes *it is enough* to consider polynomials of the type

$$V_{m_{n-1};\ell(m)}(x) = \begin{vmatrix} 1 & x_1^{m_1} & \ldots & x_1^{m_{n-1}} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n-1}^{m_1} & \ldots & x_{n-1}^{m_{n-1}} \\ 1 & x^{m_1} & \ldots & x^{m_{n-1}} \end{vmatrix} , \qquad (10)$$

with obvious meaning for the symbols used. In order to proceed we need two other concepts related to partitions (cf. [5, p. 2]).

Definition 3.2 *Given a partition λ, its* `diagram` *is the set of points $(i,j) \in \mathbb{Z}^2$ such that $1 \leq j \leq \lambda_i$.*

The **conjugate** *of a partition λ, is the partition λ' whose diagram is the transpose of the diagram of λ obtained by reflection in the main diagonal.*

These definitions allow to introduce an equivalent relation to the Jacobi-Trudi identity for Schur functions. That is, for any partition λ of n,

$$s_\lambda = \det \left(e_{\lambda'_i - i + j} \right)_{1 \leq i,j \leq n} , \qquad (11)$$

where λ' is the conjugate partition.

The following Theorem generalizes the Theorem 2.1 and its consequence, the Corollary 2.2.

Theorem 3.1 *Let $\alpha = (m_{n-1}, m_{n-2}, ..., m_1, 0)$ and $\delta = (n-1, n-2, ..., 1, 0)$ be the partitions of length n associated to the determinant (10). Let $\lambda = \alpha - \delta$ and λ' be its conjugate partition.*

$$V_{m_{n-1};\ell(m)}(x) = VDM_{n-1} \left(\prod_{i=1}^{n-1} (x - x_i) \right) P_M(x) . \qquad (12)$$

where $M = m_{n-1} - n + 1$. Moreover, setting

$$\ell_{i,j} = (-1)^{\lambda'_i - [(i-j)+1]} e_{\lambda'_i - [(i-j)+1]}$$

for $i = 1, ..., M$ and $j = 1, ..., M+1$, if $\lambda_1 \geq n$, then

$$P_M(x) = \begin{vmatrix} \ell_{1,1} & \ell_{1,2} & \cdots & \ell_{1,M+1} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \ell_{M,1} & \ell_{M,2} & \cdots & \ell_{M,M+1} \\ 1 & x & \cdots & x^M \end{vmatrix} . \qquad (13)$$

otherwise λ' has some zeros, that correspondingly reduce the degree M.

Proof. The equality (12) is simply proved by resorting to the definition of Schur functions when $x = x_n$.

Since $\lambda_1 \geq n$, then λ' has no zeros. To get the determinant (13) we observe that $P_M(x)$ is the determinant of the matrix obtained by bordering the matrix

$$\left(e_{\lambda'_i - i + j}\right)_{1 \leq i, j \leq M}, \tag{14}$$

with the column consisting of the elementary symmetric functions translated by 1 to the left starting from the first column of the matrix (14), i.e

$$\left((-1)^{\lambda'_1 - 1} e_{\lambda'_1 - 1},\ \ldots, (-1)^{\lambda'_n - M} e_{\lambda'_n - M}, 1\right)^T$$

and with the last row consisting of the vector of the powers $(1, x,\ \ldots, x^M)$.

If $\lambda_1 < n$, then λ' has $n - \lambda_1$ zeros. Therefore in the matrix (14) there are not the corresponding rows. □

Theorem 3.1 can now be used to write an algorithm for computing the coefficients of $P_M(x)$.

Algorithm 2

`Inputs:` the partitions α, δ of M and the points $x_i,\ i = 1, \ldots, M$.

1. Let $\lambda = \alpha - \delta$ and λ' its conjugate.

Build the matrix $E = \left(e_{\lambda'_i - i + j}\right)_{1 \leq i, j \leq M}$

where e_p is the pth elementary symmetric function of M points.

2. Build the matrix E_1 by bordering the matrix E: add as first column the array

$$\left((-1)^{\lambda'_1 - 1} e_{\lambda'_1 - 1},\ \ldots, (-1)^{\lambda'_n - M} e_{\lambda'_n - M}, 1\right)^T$$

and as last row $(1, x,\ \ldots, x^M)$.

`Output:` the coefficients of $P_M(x)$, which are the minors of order $M - 1$ of the matrix E_1.

References

[1] P. J. Davis. *Interpolation & Approximation.* Dover Publications, 1963.

[2] S. De Marchi. Generalized Vandermonde determinants, Toeplitz matrices and Schur functions. Rapporto di Ricerca 2/2000, Dipartimento di Matematica e Informatica. Università di Udine. Available at `www.dimi.uniud.it/~demarchi#papers`

[3] E. R. Heineman. Generalized Vandermonde determinants. Proc. Royal Soc. Edinburgh, 464–476, 1929.

[4] R. C. King. Generalized Vandermonde determinants and Schur functions. Proc. AMS **48**(1), 53–56, 1975.

[5] I. G. Macdonalds. *Symmetric Functions and Hall Polynomials.* Oxford Mathematical Monographs, Second Edition, 1995.