Numerische Mathematik 4, 117—123 (1962)

On inverses of Vandermonde and confluent Vandermonde matrices

By

WALTER GAUTSCHI *

1. Introduction

A Vandermonde matrix of order n is a matrix of the form

$$(1.1)\qquad V_n = V_n(x_1, x_2, \ldots, x_n) = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ \cdot & \cdot & \cdot & \cdot \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{pmatrix} \quad (n > 1),$$

where x_ν are real or complex numbers. By a confluence of the l-th column into the k-th column we mean the following limit operation: Replace in the l-th column x_l by $x_k + \varepsilon$ and subtract from it the k-th column; divide this new l-th column by ε and then let $\varepsilon \to 0$.

If the resulting matrix is denoted by $U_{n,kl}$ we have

$$(1.2)\qquad U_{n,kl} = \begin{pmatrix} 1 & \ldots & 1 & 0 & 1 & \ldots & 1 \\ x_1 & \ldots & x_{l-1} & 1 & x_{l+1} & \ldots & x_n \\ x_1^2 & \ldots & x_{l-1}^2 & 2x_k & x_{l+1}^2 & \ldots & x_n^2 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & \ldots & x_{l-1}^{n-1} & (n-1)\,x_k^{n-2} & x_{l+1}^{n-1} & \ldots & x_n^{n-1} \end{pmatrix}.$$

In other words, $U_{n,kl}$ is the same matrix as V_n except for the l-th column, which is the derivative of the k-th column.

A matrix that is obtained from (1.1) by one or more confluences of columns is called a confluent Vandermonde matrix. The following, for example, is a confluent Vandermonde matrix of order $2n$, obtained by confluences of the columns $n+1$ into 1, $n+2$ into 2, ..., $2n$ into n:

$$(1.3)\qquad U_{2n} = \begin{pmatrix} 1 & \ldots & 1 & 0 & \ldots & 0 \\ x_1 & \ldots & x_n & 1 & \ldots & 1 \\ \vdots & & \vdots & \vdots & & \vdots \\ x_1^{2n-1} & \ldots & x_n^{2n-1} & (2n-1)\,x_1^{2n-2} & \ldots & (2n-1)\,x_n^{2n-2} \end{pmatrix}.$$

---

* Oak Ridge National Laboratory, operated by Union Carbide Corporation for the U.S. Atomic Energy Commission, Oak Ridge, Tennessee.

The purpose of this paper is to estimate the norm of inverses of Vandermonde and confluent Vandermonde matrices. Such estimates are expected to be useful in various questions of numerical analysis. In the construction of Gauss-type quadrature formulas, for example, norm estimates of the inverse of the matrix (1.3) may be used to assess the errors in the zeros and weight factors from those in the moments.

It will be convenient to adopt the following matrix norm,

$$\|A\| = \max_{1\leq\nu\leq n} \sum_{\mu=1}^{n} |a_{\nu\mu}|, \qquad A = (a_{\nu\mu}). \tag{1.4}$$

The use of this particular norm is no real restriction since for any other norm $\|A\|_1$, one has $m\|A\| \leq \|A\|_1 \leq M\|A\|$ with positive constants m, M depending only on n, and not on A (see [3], Satz IV).

2. Preliminaries

We denote by σ_m the m-th elementary symmetric function in the n variables $x_1, x_2, \ldots, x_n$,

$$\sigma_m = \sigma_m(x_1, \ldots, x_n) = \sum x_{\nu_1} x_{\nu_2} \ldots x_{\nu_m} \qquad (1 \leq m \leq n), \qquad \sigma_0 = 1.$$

Lemma. *We have*

$$1 + |\sigma_1| + |\sigma_2| + \cdots + |\sigma_n| \leq \prod_{\nu=1}^{n} (1 + |x_\nu|), \tag{2.1}$$

where equality holds if and only if all x_ν are located on the same ray through the origin, that is, if and only if

$$x_\nu = |x_\nu| e^{i\varphi} \qquad (\nu = 1, 2, \ldots, n). \tag{2.2}$$

Proof. Let $p(x) = \prod_{\nu=1}^{n} (x - x_\nu)$. Then

$$p(x) = \sum_{m=0}^{n} (-1)^m \sigma_m x^{n-m}.$$

In particular,

$$p(-1) = (-1)^n \sum_{m=0}^{n} \sigma_m. \tag{2.3}$$

On the other hand, by definition,

$$p(-1) = (-1)^n \prod_{\nu=1}^{n} (1 + x_\nu). \tag{2.4}$$

We distinguish three cases.

Case I. All $x_\nu \geq 0$. Then all $\sigma_m \geq 0$, and from (2.3) and (2.4) we find

$$\sum_{m=0}^{n} |\sigma_m| = \sum_{m=0}^{n} \sigma_m = (-1)^n p(-1) = \prod_{\nu=1}^{n} (1 + x_\nu) = \prod_{\nu=1}^{n} (1 + |x_\nu|).$$

This proves (2.1) with equality sign.

Case II. All x_ν satisfy (2.2). Then $\sigma_m(x_1, \ldots, x_n) = e^{im\varphi}\sigma_m(|x_1|, \ldots, |x_n|)$, and

$$\sum_{m=0}^{n} |\sigma_m| = \sum_{m=0}^{n} \sigma_m(|x_1|, \ldots, |x_n|) = \prod_{\nu=1}^{n} (1+|x_\nu|)$$

by the result of Case I.

Case III. There is at least one pair of variables, say (x_1, x_2), such that $x_1 x_2 \neq 0$, $\arg x_1 \neq \arg x_2$. Then

$$|\sigma_1| = |x_1 + x_2 + \cdots + x_n| \leqq |x_1 + x_2| + |x_3| + \cdots + |x_n|$$
$$< |x_1| + |x_2| + |x_3| + \cdots + |x_n|,$$

that is,

$$|\sigma_1(x_1, \ldots, x_n)| < \sigma_1(|x_1|, \ldots, |x_n|).$$

Since also

$$|\sigma_m(x_1, \ldots, x_n)| \leqq \sigma_m(|x_1|, \ldots, |x_n|) \qquad (m > 1),$$

we find, using again the result of Case I,

$$\sum_{m=0}^{n} |\sigma_m(x_1, \ldots, x_n)| < \sum_{m=0}^{n} \sigma_m(|x_1|, \ldots, |x_n|) = \prod_{\nu=1}^{n} (1+|x_\nu|).$$

This proves (2.1) with strict inequality, and the lemma is completely proved.

Later we also use the notation σ_m^λ to denote the m-th elementary symmetric function in the $n-1$ variables x_ν with x_λ missing,

$$\sigma_m^\lambda = \sigma_m(x_1, \ldots, x_{\lambda-1}, x_{\lambda+1}, \ldots, x_n).$$

By the symmetry of σ_m we have for $\lambda < \mu$

$$\begin{aligned} \sigma_m^\lambda&(x_1, \ldots, x_{\lambda-1}, x_{\lambda+1}, \ldots, x_{\mu-1}, t, x_{\mu+1}, \ldots, x_n) \\ &= \sigma_m^\mu(x_1, \ldots, x_{\lambda-1}, t, x_{\lambda+1}, \ldots, x_{\mu-1}, x_{\mu+1}, \ldots, x_n). \end{aligned} \tag{2.5}$$

3. Inverse of Vandermonde matrix

We prove now

Theorem 1. *Let $x_\nu \neq x_\mu$ for $\nu \neq \mu$. Then, with the matrix norm defined in* (1.4), *we have*

$$\|V_n^{-1}\| \leqq \max_{1 \leqq \lambda \leqq n} \prod_{\substack{\nu=1 \\ \nu \neq \lambda}}^{n} \frac{1+|x_\nu|}{|x_\nu - x_\lambda|}. \tag{3.1}$$

If the x_ν satisfy (2.2), *then* (3.1) *is actually an equality.*

Proof. Let $V_n^{-1} = (v_{\lambda\mu})$. It is well known (see [2, p. 306], or [1]) that

$$v_{\lambda\mu} = (-1)^{\mu-1} \frac{\sigma_{n-\mu}^\lambda}{\prod_{\nu \neq \lambda} (x_\nu - x_\lambda)}. \tag{3.2}$$

Therefore,

$$\sum_{\mu=1}^{n} |v_{\lambda\mu}| = \frac{\sum_{\mu=1}^{n} |\sigma_{n-\mu}^\lambda|}{\prod_{\nu \neq \lambda} |x_\nu - x_\lambda|} \qquad (\lambda = 1, 2, \ldots, n).$$

Theorem 1 now follows immediately from the lemma in section 2.

We note that the last statement in Theorem 1 cannot be reversed, that is, if (3.1) holds with equality sign then it does not necessarily follow that all x_ν lie on the same ray through the origin. This is shown by the example $n=3$, $x_1=8$, $x_2=2$, $x_3=-1$, for which

$$V_3=\begin{pmatrix}1 & 1 & 1\\ 8 & 2 & -1\\ 64 & 4 & 1\end{pmatrix},\qquad V_3^{-1}=\begin{pmatrix}-2/54 & -1/54 & 1/54\\ 8/18 & 7/18 & -1/18\\ 16/27 & -10/27 & 1/27\end{pmatrix}.$$

Here, $\|V_3^{-1}\|=\max(4/54,\, 16/18,\, 1)=1$, and the bound on the right of (3.1) equals $\max(1/9,\, 1,\, 1)=1$, so that (3.1) is in fact an equality, even though $x_1 x_3<0$.

4. Inverses of confluent Vandermonde matrices

In this section we establish norm estimates for the inverses of the confluent matrices $U_{n,kl}$, U_{2n} defined in (1.2) and (1.3), respectively. At the same time explicit expressions are derived for the elements of $U_{n,kl}^{-1}$.

Theorem 2. *Let $x_\nu \neq x_\mu$ for $\nu\neq\mu$ $(\nu,\mu=1,2,\ldots,l-1,l+1,\ldots,n)$, and let*

$$(4.1)\qquad a_\lambda=\begin{cases}\dfrac{1+|x_k|}{|x_k-x_\lambda|} & (\lambda\neq k,l)\\[2ex] \max\left[1+|x_k|,\; 1+\left(1+|x_k|\right)\displaystyle\sum_{\substack{\nu=1\\ \nu\neq k,l}}^{n}\frac{1}{|x_\nu-x_k|}\right] & (\lambda=k).\end{cases}$$

Then, with the matrix norm defined in (1.4), *we have*

$$(4.2)\qquad \|U_{n,kl}^{-1}\|\leq\max_{\substack{1\leq\lambda\leq n\\ \lambda\neq l}} a_\lambda \prod_{\substack{\nu=1\\ \nu\neq\lambda,l}}^{n}\frac{1+|x_\nu|}{|x_\nu-x_\lambda|}.$$

Proof. Assume for the sake of definiteness that $k<l$. Let us introduce the "perturbed" Vandermonde matrix

$$V_{n,kl}(\varepsilon)=V_n(x_1,\ldots,x_{l-1},x_k+\varepsilon,x_{l+1},\ldots,x_n),$$

and the auxiliary matrix

$$(4.3)\qquad E_{kl}(\varepsilon)=\begin{array}{c@{}l}\begin{pmatrix}1 & \ldots & 0 & \ldots & 0 & \ldots & 0\\ \vdots & & \vdots & & \vdots & & \vdots\\ 0 & \ldots & 1 & \ldots & -\varepsilon^{-1} & \ldots & 0\\ \vdots & & \vdots & & \vdots & & \vdots\\ 0 & \ldots & 0 & \ldots & \varepsilon^{-1} & \ldots & 0\\ \vdots & & \vdots & & \vdots & & \vdots\\ 0 & \ldots & 0 & \ldots & 0 & \ldots & 1\end{pmatrix} & \begin{array}{l}\\ \\ k\text{-th row}\\ \\ l\text{-th row}\\ \\ \\ \end{array}\\ \begin{array}{ccc}\qquad & k\text{-th column} & l\text{-th column}\end{array} & \end{array}.$$

Then it is not difficult to see, that by definition of confluence,

$$U_{n,kl}=\lim_{\varepsilon\to 0}V_{n,kl}(\varepsilon)\,E_{kl}(\varepsilon).$$

From this we get

$$U_{n,kl}^{-1} = \lim_{\varepsilon\to 0} E_{kl}^{-1}(\varepsilon)\, V_{n,kl}^{-1}(\varepsilon), \tag{4.4}$$

provided that the limit on the right-hand side exists.

Inverting (4.3) we have

$$E_{kl}^{-1}(\varepsilon) = \begin{array}{cc} \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & \varepsilon & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} & \begin{matrix} \\ \\ k\text{-th row} \\ \\ l\text{-th row} \\ \\ \\ \end{matrix} \\ \begin{matrix} \qquad\qquad k\text{-th} & l\text{-th} \\ \qquad\qquad \text{column} & \text{column} \end{matrix} & \end{array}.$$

Therefore, if $V_{n,kl}^{-1}(\varepsilon) = [v_{\lambda\mu}(\varepsilon)]$, we find

$$E_{kl}^{-1}(\varepsilon)\, V_{n,kl}^{-1}(\varepsilon) = \begin{pmatrix} v_{11}(\varepsilon) & \dots & v_{1n}(\varepsilon) \\ \vdots & & \vdots \\ v_{k1}(\varepsilon) + v_{l1}(\varepsilon) & \dots & v_{kn}(\varepsilon) + v_{ln}(\varepsilon) \\ \vdots & & \vdots \\ \varepsilon\, v_{l1}(\varepsilon) & \dots & \varepsilon\, v_{ln}(\varepsilon) \\ \vdots & & \vdots \\ v_{n1}(\varepsilon) & \dots & v_{nn}(\varepsilon) \end{pmatrix}.$$

Since $V_{n,kl}$ is a Vandermonde matrix, the elements of its inverse are given by (3.2), that is

$$v_{\lambda\mu}(\varepsilon) = (-1)^{\mu-1} \frac{\sigma_{n-\mu}^{\lambda}}{\prod\limits_{\nu\neq\lambda} (x_\nu - x_\lambda)}. \tag{4.5}$$

It is understood here, that x_l, wherever it occurs, is to be replaced by $x_k + \varepsilon$. If $\lambda \neq k, l$ the expression in (4.5) has a well defined limit, as $\varepsilon \to 0$, namely

$$\lim_{\varepsilon\to 0} v_{\lambda\mu}(\varepsilon) = (-1)^{\mu-1} \frac{\sigma_{n-\mu}^{\lambda}(x_1, \dots, x_{l-1}, x_k, x_{l+1}, \dots, x_n)}{(x_k - x_\lambda) \prod\limits_{\nu\neq\lambda, l} (x_\nu - x_\lambda)} \qquad (\lambda \neq k, l). \tag{4.6}$$

If $\lambda = k$, we have, using (2.5), with $\lambda = k$, $\mu = l$, $t = x_k + \varepsilon$,

$$\begin{aligned} v_{k\mu}(\varepsilon) &= (-1)^{\mu-1} \frac{\sigma_{n-\mu}^{k}(x_1, \dots, x_{l-1}, x_k + \varepsilon, x_{l+1}, \dots, x_n)}{\varepsilon \prod\limits_{\nu\neq k, l} (x_\nu - x_k)} \\ &= \frac{(-1)^{\mu-1}}{\varepsilon} \frac{\sigma_{n-\mu}^{l}(x_1, \dots, x_{k-1}, x_k + \varepsilon, x_{k+1}, \dots, x_n)}{\prod\limits_{\nu\neq k, l} (x_\nu - x_k)}. \end{aligned} \tag{4.7}$$

If, finally, $\lambda = l$ then

$$v_{l\mu}(\varepsilon) = \frac{(-1)^{\mu}}{\varepsilon} \frac{\sigma_{n-\mu}^{l}(x_1, \dots, x_{k-1}, x_k, x_{k+1}, \dots, x_n)}{\prod\limits_{\nu\neq k, l} (x_\nu - x_k - \varepsilon)}. \tag{4.8}$$

The sum of the two expressions in (4.7) and (4.8) is seen to have the form $(-1)^{\mu-1}\varepsilon^{-1}[\sigma(x+\varepsilon)\pi^{-1}(x)-\sigma(x)\pi^{-1}(x+\varepsilon)]$ where σ, π stand for the numerator and denominator functions, both considered as functions of $x=x_k$. Since

$$\frac{\sigma(x+\varepsilon)}{\pi(x)}-\frac{\sigma(x)}{\pi(x+\varepsilon)}=\varepsilon\frac{\frac{d}{dx}[\sigma(x)\,\pi(x)]}{\pi^2(x)}+o(\varepsilon)\qquad(\varepsilon\to 0)$$

we obtain

$$\text{(4.9)}\qquad \lim_{\varepsilon\to 0}[v_{k\mu}(\varepsilon)+v_{l\mu}(\varepsilon)]=(-1)^{\mu-1}\frac{\frac{\partial}{\partial x_k}\left[\sigma^l_{n-\mu}\prod_{\nu\neq k,l}(x_\nu-x_k)\right]}{\prod_{\nu\neq k,l}(x_\nu-x_k)^2}.$$

Let us carry out the differentiation in the numerator. We first observe that

$$\frac{\partial}{\partial x_k}\sigma^l_{n-\mu}=\sigma^{l,k}_{n-\mu-1}\qquad(\mu=1,2,\ldots,n),\qquad \sigma^{l,k}_{-1}=0,$$

where $\sigma^{l,k}_m$ denotes the m-th elementary symmetric function in the $n-2$ variables x_ν with both x_l and x_k missing. Next we note that

$$\frac{\frac{\partial}{\partial x_k}\prod_{\nu\neq k,l}(x_\nu-x_k)}{\prod_{\nu\neq k,l}(x_\nu-x_k)}=-\sum_{\nu\neq k,l}\frac{1}{x_\nu-x_k}.$$

Therefore, we obtain from (4.9)

$$\text{(4.10)}\qquad \lim_{\varepsilon\to 0}[v_{k\mu}(\varepsilon)+v_{l\mu}(\varepsilon)]=\frac{(-1)^{\mu-1}}{\prod_{\nu\neq k,l}(x_\nu-x_k)}\left\{\sigma^{l,k}_{n-\mu-1}-\sigma^l_{n-\mu}\sum_{\nu\neq k,l}\frac{1}{x_\nu-x_k}\right\}.$$

Finally, from (4.8) we see that

$$\text{(4.11)}\qquad \lim_{\varepsilon\to 0}\varepsilon\, v_{l\mu}(\varepsilon)=(-1)^{\mu}\frac{\sigma^l_{n-\mu}}{\prod_{\nu\neq k,l}(x_\nu-x_k)}.$$

The relations (4.6), (4.10) and (4.11) now show not only that the limiting matrix in (4.4), and thus $U^{-1}_{n,kl}$, exists, but they also give explicit expressions for the elements $u_{\lambda\mu}$ of $U^{-1}_{n,kl}$. From these, and from the lemma in section 2 we conclude

$$\sum_{\mu=1}^{n}|u_{\lambda\mu}|\leqq\frac{1+|x_k|}{|x_k-x_\lambda|}\prod_{\nu\neq\lambda,l}\frac{1+|x_\nu|}{|x_\nu-x_\lambda|}\qquad(\lambda\neq k,l),$$

$$\sum_{\mu=1}^{n}|u_{k\mu}|\leqq\left\{1+(1+|x_k|)\sum_{\nu\neq k,l}\frac{1}{|x_\nu-x_k|}\right\}\prod_{\nu\neq k,l}\frac{1+|x_\nu|}{|x_\nu-x_k|},$$

$$\sum_{\mu=1}^{n}|u_{l\mu}|\leqq(1+|x_k|)\prod_{\nu\neq k,l}\frac{1+|x_\nu|}{|x_\nu-x_k|},$$

which is equivalent to (4.1), (4.2). Theorem 2 is proved.

The argument in the proof of Theorem 2 can be applied repeatedly to deal with matrices that are derived from a Vandermonde matrix by more than one confluence of columns. One so obtains, for example, the following

Theorem 3. *Let* $x_\nu \neq x_\mu$ *for* $\nu \neq \mu$ $(\nu, \mu = 1, 2, \ldots, n)$, *and let*

$$b_\lambda = \max\left[1 + |x_\lambda|,\ 1 + 2(1 + |x_\lambda|) \sum_{\substack{\nu=1 \\ \nu \neq \lambda}}^{n} \frac{1}{|x_\nu - x_\lambda|}\right]. \tag{4.12}$$

Then, with the matrix norm defined in (1.4), *we have*

$$\|U_{2n}^{-1}\| \leq \max_{1 \leq \lambda \leq n} b_\lambda \left(\prod_{\substack{\nu=1 \\ \nu \neq \lambda}}^{n} \frac{1 + |x_\nu|}{|x_\nu - x_\lambda|}\right)^2. \tag{4.13}$$

References

[1] MACON, N., and A. SPITZBART: Inverses of Vandermonde matrices. Amer. Math. Monthly **65**, 95—100 (1958).

[2] MUIR, TH.: The theory of determinants, vol. I. Reprinted in Dover Publications, New York, 1960.

[3] OSTROWSKI, A.: Über Normen von Matrizen. Math. Z. **63**, 2—18 (1955).

Oak Ridge National Laboratory
Mathematics Panel
Post Office Box X
Oak Ridge, Tennessee

(Received October 18, 1961)