What Is a Vandermonde Matrix?

Nicholas J. Higham*

June 15, 2021

A Vandermonde matrix is defined in terms of scalars x_1, x_2, $\dots$, $x_n \in \mathbb{C}$ by

$$V = V(x_1, x_2, \dots, x_n) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{bmatrix} \in \mathbb{C}^{n\times n}.$$

The x_i are called *points* or *nodes*. Note that while we have indexed the nodes from 1, they are usually indexed from 0 in papers concerned with algorithms for solving Vandermonde systems.

Vandermonde matrices arise in polynomial interpolation. Suppose we wish to find a polynomial $p_{n-1}(x) = a_n x^{n-1} + a_{n-1}x^{n-2} + \dots + a_1$ of degree at most $n-1$ that interpolates to the data $(x_i, f_i)_{i=1}^n$, that is, $p_{n-1}(x_i) = f_i$, $i = 1\colon n$. These equations are equivalent to

$$V^T a = f \quad \text{(dual)},$$

where $a = [a_1, a_2, \dots, a_n]^T$ is the vector of coefficients. This is known as the *dual problem*. We know from polynomial interpolation theory that there is a unique interpolant if the x_i are distinct, so this is the condition for V to be nonsingular.

The problem

$$Vy = b \quad \text{(primal)}$$

is called the *primal problem*, and it arises when we determine the weights for a quadrature rule: given moments b_i find weights y_i such that $\sum_{j=1}^n y_j x_j^{i-1} = b_i$, $i = 1\colon n$.

Determinant

The determinant of V is a function of the n points x_i. If $x_i = x_j$ for some $i \ne j$ then V has identical ith and jth columns, so is singular. Hence the determinant must have a factor $x_i - x_j$. Consequently, we have

$$\det(V(x_1, x_2, \dots, x_n)) = c \prod_{\substack{i,j=1\\ i>j}}^{n} (x_i - x_j),$$

*Department of Mathematics, University of Manchester, Manchester, M13 9PL, UK (nick.higham@manchester.ac.uk).

where, since both sides have degree $n(n-1)/2$ in the x_i, c is a constant. But $\det(V)$ contains a term $x_2 x_3^2 \dots x_n^{n-1}$ (from the main diagonal), so $c = 1$. Hence

$$\det(V) = \prod_{\substack{i,j=1 \\ i>j}}^{n} (x_i - x_j). \qquad (1)$$

This formula confirms that V is nonsingular precisely when the x_i are distinct.

Inverse

Now assume that V is nonsingular and let $V^{-1} = W = (w_{ij})_{i,j=1}^{n}$. Equating elements in the ith row of $WV = I$ gives

$$\sum_{j=1}^{n} w_{ij} x_k^{j-1} = \delta_{ik}, \quad k = 1\colon n,$$

where δ_{ij} is the Kronecker delta (equal to 1 if $i = j$ and 0 otherwise). These equations say that the polynomial $\sum_{j=1}^{n} w_{ij} x^{j-1}$ takes the value 1 at $x = x_i$ and 0 at $x = x_k$, $k \neq i$. It is not hard to see that this polynomial is the Lagrange basis polynomial:

$$\sum_{j=1}^{n} w_{ij} x^{j-1} = \prod_{\substack{k=1 \\ k\neq i}}^{n} \left(\frac{x - x_k}{x_i - x_k} \right) =: \ell_i(x). \qquad (2)$$

We deduce that

$$w_{ij} = \frac{(-1)^{n-j} \sigma_{n-j}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)}{\displaystyle\prod_{\substack{k=1 \\ k\neq i}}^{n} (x_i - x_k)}, \qquad (3)$$

where $\sigma_k(y_1, \dots, y_n)$ denotes the sum of all distinct products of k of the arguments $y_1, \dots, y_n$ (that is, σ_k is the kth elementary symmetric function).

From (1) and (3) we see that if the x_i are real and positive and arranged in increasing order $0 < x_1 < x_2 < \dots < x_n$ then $\det(V) > 0$ and V^{-1} has a checkerboard sign pattern: the (i, j) element has sign $(-1)^{i+j}$.

Note that summing (2) over i gives

$$\sum_{j=1}^{n} x^{j-1} \sum_{i=1}^{n} w_{ij} = \sum_{i=1}^{n} \ell_i(x) = 1,$$

where the second equality follows from the fact that $\sum_{i=1}^{n} \ell_i(x)$ is a degree $n-1$ polynomial that takes the value 1 at the n distinct points x_i. Hence

$$\sum_{i=1}^{n} w_{ij} = \delta_{j1},$$

so the elements in the jth column of the inverse sum to 1 for $j = 1$ and 0 for $j \geq 2$.

Example

To illustrate the formulas above, here is an example, with $x_i = (i-1)/(n-1)$ and $n = 5$:

$$V = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} & 1 \\ 0 & \frac{1}{16} & \frac{1}{4} & \frac{9}{16} & 1 \\ 0 & \frac{1}{64} & \frac{1}{8} & \frac{27}{64} & 1 \\ 0 & \frac{1}{256} & \frac{1}{16} & \frac{81}{256} & 1 \end{bmatrix}, \quad V^{-1} = \begin{bmatrix} 1 & -\frac{25}{3} & \frac{70}{3} & -\frac{80}{3} & \frac{32}{3} \\ 0 & 16 & -\frac{208}{3} & 96 & -\frac{128}{3} \\ 0 & -12 & 76 & -128 & 64 \\ 0 & \frac{16}{3} & -\frac{112}{3} & \frac{224}{3} & -\frac{128}{3} \\ 0 & -1 & \frac{22}{3} & -16 & \frac{32}{3} \end{bmatrix},$$

for which $\det(V) = 9/32768$.

Conditioning

Vandermonde matrices are notorious for being ill conditioned. The ill conditioning stems from the monomials being a poor basis for the polynomials on the real line. For arbitrary distinct points x_i, Gautschi showed that $V_n = V(x_1, x_2, \dots, x_n)$ satisfies

$$\max_i \prod_{j \neq i} \frac{\max(1, |x_j|)}{|x_i - x_j|} \leq \|V_n^{-1}\|_\infty \leq \max_i \prod_{j \neq i} \frac{1 + |x_j|}{|x_i - x_j|},$$

with equality on the right when $x_j = |x_j| e^{\mathrm{i}\theta}$ for all j with a fixed θ (in particular, when $x_j \geq 0$ for all j). Note that the upper and lower bounds differ by at most a factor 2^{n-1}. It is also known that for any set of real points x_i,

$$\kappa_2(V_n) \geq \left(\frac{2}{n}\right)^{1/2} (1 + \sqrt{2})^{n-2}$$

and that for $x_i = 1/i$ we have $\kappa_\infty(V_n) > n^{n+1}$, where the lower bound is an extremely fast growing function of the dimension!

These exponential lower bounds are alarming, but they do not necessarily rule out the use of Vandermonde matrices in practice. One of the reasons is that there are specialized algorithms for solving Vandermonde systems whose accuracy is not dependent on the condition number κ, and which in some cases can be proved to be highly accurate. The first such algorithm is an $O(n^2)$ operation algorithm for solving $V_n y = b$ of Björck and Pereyra (1970). There is now a long list of generalizations of this algorithm in various directions, including for confluent Vandermonde-like matrices (Higham, 1990), as well as for more specialized problems (Demmel and Koev, 2005) and more general ones (Bella et al., 2009). Another important observation is that the exponential lower bounds are for real nodes. For complex nodes V_n can be much better conditioned. Indeed when the x_i are the roots of unity, $V_n/\sqrt{n}$ is the unitary Fourier matrix and so V_n is perfectly conditioned.

Generalizations

Two ways in which Vandermonde matrices have been generalized are by allowing confluency of the points x_i and by replacing the monomials by other polynomials. Confluency arises when the x_i are not distinct. If we assume that equal x_i are contiguous then a confluent Vandermonde matrix is obtained by "differentiating" the previous column for each of the repeated points. For example, with points x_1, x_1, x_1, x_2, x_2 we obtain

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ x_1 & 1 & 0 & x_2 & 1 \\ x_1^2 & 2x_1 & 2 & x_2^2 & 2x_2 \\ x_1^3 & 3x_1^2 & 6x_1 & x_2^3 & 3x_2^2 \\ x_1^4 & 4x_1^3 & 12x_1^2 & x_2^4 & 4x_2^3 \end{bmatrix}. \qquad (4)$$

The transpose of a confluent Vandermonde matrix arises in Hermite interpolation; it is nonsingular if the points corresponding to the "nonconfluent columns" are distinct (that is, if $x_1 \neq x_2$ in the case of (4)).

A *Vandermonde-like matrix* is defined in terms of a set of polynomials $\{p_i(x)\}_{i=0}^n$ with p_i having degree i:

$$\begin{bmatrix} p_0(x_1) & p_0(x_2) & \dots & p_0(x_n) \\ p_1(x_1) & p_1(x_2) & \dots & p_1(x_n) \\ \vdots & \vdots & \dots & \vdots \\ p_{n-1}(x_1) & p_{n-1}(x_2) & \dots & p_{n-1}(x_n) \end{bmatrix}.$$

Of most interest are polynomials that satisfy a three-term recurrence, in particular, orthogonal polynomials. Such matrices can be much better conditioned than general Vandermonde matrices.

Notes

Algorithms for solving confluent Vandermonde-like systems and their rounding error analysis are described in the chapter "Vandermonde systems" of Higham (2002).

Gautschi has written many papers on the conditioning of Vandermonde matrices, beginning in 1962. We mention just his most recent paper on this topic: Gautschi (2011).

References

This is a minimal set of references, which contain further useful references within.

- T. Bella, Y. Eidelman. I. Gohberg. I. Koltracht, and V. Olshevsky, A Fast Björck–Pereyra-Type Algorithm for Solving Hessenberg-Quasiseparable-Vandermonde Systems, SIAM J. Matrix Anal. Appl. 31(2), 790–815, 2009.
- James W. Demmel and Plamen Koev, The Accurate and Efficient Solution of a Totally Positive Generalized Vandermonde Linear System, SIAM J. Matrix Anal. Appl. 2791, 142–152, 2005.
- Walter Gautschi, Optimally Scaled and Optimally Conditioned Vandermonde and Vandermonde-like matrices, BIT 51, 103–125, 2011.
- Nicholas J. Higham, Accuracy and Stability of Numerical Algorithms, second edition, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.
- Nicholas J. Higham, Stability Analysis of Algorithms for Solving Confluent Vandermonde-like Systems, SIAM J. Matrix Anal. Appl. 11, 23–41, 1990.

Related Blog Posts

- What Is a Condition Number? (2020)

This article is part of the “What Is” series, available from https://nhigham.com/category/what-is and in PDF form from the GitHub repository https://github.com/higham/what-is.