On the Inversion of Vandermonde Matrices

Shui-Hung Hou
Department of Applied Mathematics
Hong Kong Polytechnic University
mahoush@polyu.edu.hk

Abstract

A novel and simple recursive algorithm for inverting Vandermonde matrix and its generalized form is presented. The algorithm is suitable for classroom use in both numerical as well as symbolic computation.

1 Introduction

The importance of the Vandermonde matrix is well known. Inversion of this matrix is necessary in many areas of applications such as polynomial interpolation [4, 10], digital signal processing [2], and control theory [6], to mention a few. See also for example Klinger [8], Kalman [7]. However, an explicit recursive formula for the inversion of Vandermonde matrices seems unavailable in most linear algebra textbooks.

The purpose of this paper is to present a novel and simple recursive algorithm for inverting Vandermonde matrix, as well as its generalized (or confluent) form, in a way more readily accessible for use in classroom and suitable for both numerical as well as symbolic computation.

2 Preliminaries and notations

Let m be a nonnegative integer. For the sequence $1, (s-\lambda), \ldots, (s-\lambda)^{m-1}$ of polynomials we write $\mathbf{s}(\lambda, m) = [1, (s-\lambda), \ldots, (s-\lambda)^{m-1}]^T$. In particular,

$$\mathbf{s}(0, m) = [1, s, \ldots, s^{m-1}]^T.$$

Let $\lambda_1, \lambda_2, \ldots, \lambda_r$ be given distinct zeros of the polynomial

$$p(s) = (s-\lambda_1)^{n_1} \cdots (s-\lambda_r)^{n_r}$$

with $n_1 + \ldots + n_r = n$. The generalized (or confluent) Vandermonde matrix related to the zeros of $p(s)$ is known to be

$$V = [V_1 V_2 \cdots V_r], \qquad (1)$$

where the block matrix $V_k = V(\lambda_k, n_k)$ is of order $n \times n_k$, having elements $V(\lambda_k, n_k)_{ij} = \binom{i-1}{j-1}\lambda_k^{i-j}$ for $i \geq j$ and zero otherwise ($k = 1, 2, \ldots, r$; $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, n_k$). More specifically, V_k is the $n \times n_k$ matrix of coefficients that appears in the truncated Taylor expansion at λ_k, modulo $(s - \lambda_k)^{n_k}$, of $\mathbf{s}(0, n)$. That is,

$$\mathbf{s}(0, n) = V(\lambda_k, n_k)\mathbf{s}(\lambda_k, n_k) \mod (s - \lambda_k)^{n_k}.$$

In the case the zeros $\lambda_1, \ldots, \lambda_r$ of $p(s)$ are simple, we have the usual Vandermonde matrix, namely,

$$V = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_r \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n_1-1} & \lambda_2^{n_2-1} & \cdots & \lambda_r^{n_r-1} \end{bmatrix}.$$

It will be shown that the inverse of the generalized Vandermonde matrix V in (1) has a form

$$V^{-1} = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_r \end{bmatrix},$$

where each block matrix W_k is of order $n_k \times n$, and may be computed by means of a recursive procedure.

Generally, the inverse of the usual Vandermonde matrix [3], as well as the inverse of the generalized Vandermonde matrix [9] are based on using interpolation polynomials.

Our approach is based on using the Leverrier-Faddeev algorithm [1, 5, 10], which states that the resolvent of a given $n \times n$ matrix A is given by

$$(sI - A)^{-1} = \frac{B_1 s^{n-1} + B_2 s^{n-2} + \cdots + B_n}{s^n + a_1 s^{n-1} + \cdots + a_n}, \qquad (2)$$

where $\det(sI - A) = s^n + a_1 s^{n-1} + \cdots + a_n$ is the characteristic polynomial of the matrix A, and all the B_j matrices are of order $n \times n$, satisfying

$$\begin{array}{ll} B_1 = I, & a_1 = -\frac{1}{1}\operatorname{tr}(AB_1), \\ B_2 = AB_1 + a_1 I, & a_2 = -\frac{1}{2}\operatorname{tr}(AB_2), \\ \vdots & \vdots \\ B_n = AB_{n-1} + a_{n-1} I, & a_n = -\frac{1}{n}\operatorname{tr}(AB_n) \end{array} \tag{3}$$

with $0 = AB_n + a_n I$ terminating as a check of computation. Here tr stands for the trace of a matrix.

3 Main result

Let $J = \operatorname{diag}(J_1, \ldots, J_r)$ be the block diagonal matrix, where

$$J_k = J(\lambda_k, n_k) = \begin{bmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & & \vdots \\ 0 & & \ddots & \ddots & 0 \\ \vdots & & & \lambda_k & 1 \\ 0 & \cdots & 0 & 0 & \lambda_k \end{bmatrix}$$

is the $n_k \times n_k$ Jordan block with eigenvalue λ_k. Then J has characteristic polynomial $\det(sI - J) = (s - \lambda_1)^{n_1} \cdots (s - \lambda_r)^{n_r} = p(s)$.

Substituting $A = J$ in equations (2) and (3) of the Leverrier-Faddeev algorithm, we see immediately that

$$p(s)(sI - J)^{-1} = B_1 s^{n-1} + B_2 s^{n-2} + \cdots + B_n, \tag{4}$$

where

$$\begin{array}{rcl} B_1 & = & I, \\ B_2 & = & JB_1 + a_1 I, \\ & \cdots & \\ B_n & = & JB_{n-1} + a_{n-1} I, \\ 0 & = & JB_n + a_n I. \end{array}$$

$J = \operatorname{diag}(J_1, \ldots, J_r)$ being block diagonal, so are all the B_j matrices. In fact,

$$B_j = \operatorname{diag}(B_{j,1}, B_{j,2}, \ldots, B_{j,r}), \quad j = 1, 2, \ldots, n,$$

and each block matrix $B_{j,k}$ is of order $n_k \times n_k$, satisfying

$$\begin{array}{rcl} B_{1,k} & = & I_k, \\ B_{2,k} & = & J_k B_{1,k} + a_1 I_k, \\ & \cdots & \\ B_{n,k} & = & J_k B_{n-1,k} + a_{n-1} I_k, \\ 0 & = & J_k B_{n,k} + a_n I_k, \end{array} \tag{5}$$

where I_k is the $n_k \times n_k$ identity matrix.
Let us now put

$$p_k(s) = \frac{p(s)}{(s-\lambda_k)^{n_k}}, \qquad k = 1, \ldots, r.$$

Define also the n_k-dimensional column vector $\theta_k = [0, \cdots, 0, 1]^T$, and write $\theta = [\theta_1^T, \cdots, \theta_r^T]^T$.
If we postmultiply both sides of equation (4) by the column vector θ, we easily get

$$\begin{bmatrix} p_1(s)\mathbf{s}(\lambda_1, n_1) \\ \vdots \\ p_r(s)\mathbf{s}(\lambda_r, n_r) \end{bmatrix} = \begin{bmatrix} H_1 \\ \vdots \\ H_r \end{bmatrix} \mathbf{s}(0, n). \tag{6}$$

Each H_k is of the form

$$H_k = \begin{bmatrix} B_{n,k}\theta_k & \cdots & B_{1,k}\theta_k \end{bmatrix} \tag{7}$$

and has order $n_k \times n$.
Comparing in turn for $k = 1, 2, \ldots, r$ the truncated Taylor expansions at λ_k, modulo $(s-\lambda_k)^{n_k}$, of both sides in (6) and putting these results together, we get

$$\text{diag}(P_1, \ldots, P_r) = \begin{bmatrix} H_1 \\ \vdots \\ H_r \end{bmatrix} \begin{bmatrix} V_1 & \cdots & V_r \end{bmatrix},$$

where each block P_k is a $n_k \times n_k$ upper triangular matrix given by

$$P_k = p_k(J_k) = \sum_{j=0}^{n_k-1} \frac{p_k^{(j)}(\lambda_k)}{j!}(N_k)^j.$$

It is noted here that $N_k = J(0, n_k) = J_k - \lambda_k I_k$ is nilpotent of order n_k.
If we can show that each P_k is invertible, then

$$V^{-1} \quad = \quad \begin{bmatrix} P_1^{-1}H_1 \\ \vdots \\ P_r^{-1}H_r \end{bmatrix}. \tag{8}$$

To this end we require the following lemma which is an easy consequence of the partial fraction expansion of $1/p(s)$ and the fact that N_k is nilpotent.

Lemma 1 *Let there be given the partial fraction expansion*

$$\frac{1}{p(s)} = \sum_{k=1}^{r} \left(\frac{K_{k,n_k}}{(s-\lambda_k)^{n_k}} + \frac{K_{k,n_k-1}}{(s-\lambda_k)^{n_k-1}} + \cdots + \frac{K_{k,1}}{s-\lambda_k} \right).$$

Then for $k = 1, 2, \ldots, r$

$$P_k^{-1} = \sum_{j=0}^{n_k-1} K_{k,j}(N_k)^j = \mathcal{K}_k(J_k),$$

where the polynomial $\mathcal{K}_k(s)$ *is given by*

$$\mathcal{K}_k(s) = K_{k,n_k} + K_{k,n_k-1}(s-\lambda_k) + \cdots + K_{k,1}(s-\lambda_k)^{n_k-1}.$$

Putting the above results together with equations (7) and (8), we are now ready to state our main result:

Theorem 1 *The inverse of* $V = [V_1 V_2 \ldots V_r]$ *related to the distinct zeros* $\lambda_1, \ldots, \lambda_r$ *of* $p(s)$ *is given by*

$$V^{-1} = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_r \end{bmatrix},$$

where each block matrix

$$W_k = W(\lambda_k, n_k) = \begin{bmatrix} \mathcal{K}_k(J_k)B_{n,k}\theta_k & \mathcal{K}_k(J_k)B_{n-1,k}\theta_k & \cdots & \mathcal{K}_k(J_k)B_{1,k}\theta_k \end{bmatrix}$$

is of order $n_k \times n$.

Taking into account of (5), we find that $\mathcal{K}_k(J_k)B_{j,k} = B_{j,k}\mathcal{K}_k(J_k), j = 1, 2, \ldots, n$, so that $B_{1,k}\mathcal{K}_k(J_k) = \mathcal{K}_k(J_k)$, and for $j = 2, \ldots, n$

$$\begin{aligned} B_{j,k}\mathcal{K}_k(J_k) &= J_k B_{j-1,k}\mathcal{K}_k(J_k) + a_{j-1}\mathcal{K}_k(J_k) \\ &= (\lambda_k I_k + N_k)B_{j-1,k}\mathcal{K}_k(J_k) + a_{j-1}\mathcal{K}_k(J_k). \end{aligned}$$

Moreover, $(\lambda_k I_k + N_k)B_{n,k}\mathcal{K}_k(J_k) + a_n\mathcal{K}_k(J_k) = 0$.

4 The Algorithm

Based on the results obtained in the last section, we are now ready to give a recursive algorithm for inverting generalized Vandermonde matrix.

The Algorithm:

Let $\lambda_1, \lambda_2, \ldots, \lambda_r$ be distinct zeros of the polynomial

$$\begin{aligned} p(s) &= (s-\lambda_1)^{n_1} \cdots (s-\lambda_r)^{n_r} \\ &= s^n + a_1 s^{n-1} + \cdots + a_n \end{aligned}$$

given together with the partial fraction expansion of

$$\frac{1}{p(s)} = \sum_{k=1}^{r} \left(\frac{K_{k,n_k}}{(s-\lambda_k)^{n_k}} + \frac{K_{k,n_k-1}}{(s-\lambda_k)^{n_k-1}} + \cdots + \frac{K_{k,1}}{s-\lambda_k} \right).$$

For each $k \in \{1, 2, \ldots, r\}$, compute recursively polynomials $h_1, h_2, \ldots, h_n$ of degree at most $n_k - 1$ by means of the following scheme:

$$\begin{aligned} h_1(s) &= K_{k,n_k} + sK_{k,n_k-1} + \cdots + s^{n_k-1}K_{k,1}, \\ h_2(s) &= (\lambda_k + s)h_1(s) + a_1 h_1(s) \mod s^{n_k}, \\ h_3(s) &= (\lambda_k + s)h_2(s) + a_2 h_1(s) \mod s^{n_k}, \\ &\vdots \\ h_n(s) &= (\lambda_k + s)h_{n-1}(s) + a_{n-1} h_1(s) \mod s^{n_k}, \end{aligned}$$

terminating at

$$0 = (\lambda_k + s)h_n(s) + a_n h_1(s) \mod s^{n_k}.$$

Obtain a block matrix $W_k = W(\lambda_k, n_k)$ of order $n_k \times n$ via the equality

$$\left[\begin{array}{cccc} s^{n_k-1} & s^{n_k-2} & \cdots & 1 \end{array} \right] W(\lambda_k, n_k) = \left[\begin{array}{cccc} h_n & h_{n-1} & \cdots & h_1 \end{array} \right].$$

The inverse of the generalized Vandermonde matrix V related to the distinct zeros $\lambda_1, \lambda_2, \ldots, \lambda_r$ of $p(s)$ may then be given by

$$\begin{aligned} V^{-1} &= [V(\lambda_1, n_1)V(\lambda_2, n_2)\cdots V(\lambda_r, n_r)]^{-1} \\ &= \left[\begin{array}{c} W(\lambda_1, n_1) \\ W(\lambda_2, n_2) \\ \vdots \\ W(\lambda_r, n_r) \end{array} \right]. \end{aligned}$$

Let us now give some supplementary remarks on the above algorithm.

(i) A check on the accuracy of the computation of polynomials $h_1, \ldots, h_n$ is provided by the last polynomial $(\lambda_k + s)h_n(s) + a_n h_1(s)$, which should result identically in the zero polynomial 0 when modulo s^{n_k} is performed.

(ii) The coefficients $a_1, a_2, \ldots, a_n$ of the polynomial $p(s)$ may be recursively computed using (3) with $A = \mathrm{diag}(\underbrace{\lambda_1, \ldots, \lambda_1}_{n_1}, \ldots, \underbrace{\lambda_r, \ldots, \lambda_r}_{n_r})$.

(iii) The partial fraction coefficients $K_{k,n_k}, K_{k,n_k-1}, \ldots, K_{k,1}$ used in the construction of the starting polynomial $h_1(s)$ may be obtained by expanding

$$\frac{1}{p_k(s)} = \sum_{j=0}^{n_k-1} K_{k,n_k-j}(s-\lambda_k)^j + \cdots$$

in powers of $(s - \lambda_k)$. They may also be recursively computed using the following scheme:

$$\begin{aligned} K_{k,n_k} &= 1/p_{k,0}\ , \\ K_{k,n_k-j} &= -\frac{\sum_{i=1}^{j} p_{k,i} K_{k,n_k-j+i}}{p_{k,0}} \quad (j = 1, \ldots, n_k - 1), \end{aligned}$$

where $p_k(s) = \sum_{j=0}^{n-n_k} p_{k,j}(s-\lambda_k)^j$.

5 Illustrative example

The following example will serve to illustrate the recursive algorithm presented above. Let the generalized Vandermonde matrix V in (1) be given by

$$V = \begin{bmatrix} 1 & 0 & 0 & 1 \\ \lambda_1 & 1 & 0 & \lambda_2 \\ \lambda_1^2 & 2\lambda_1 & 1 & \lambda_2^2 \\ \lambda_1^3 & 3\lambda_1^2 & 3\lambda_1 & \lambda_2^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ -2 & 1 & 0 & 3 \\ 4 & -4 & 1 & 9 \\ -8 & 12 & -6 & 27 \end{bmatrix},$$

for which $\lambda_1 = -2, n_1 = 3$, and $\lambda_2 = 3, n_2 = 1$. The coefficients of the polynomial $p(s) = (s+2)^3(s-3)$ are given by $a_1 = 3, a_2 = -6, a_3 = -28,$

and $a_4 = -24$. It is easy to determine the partial fraction expansion of $1/p(s)$ to be

$$\frac{1}{(s+2)^3(s-3)} = \frac{-\frac{1}{5}}{(s+2)^3} + \frac{-\frac{1}{25}}{(s+2)^2} + \frac{-\frac{1}{125}}{s+2} + \frac{\frac{1}{125}}{s-3}.$$

Let us consider first the case $\lambda_1 = -2$. Clearly

$$h_1(s) = -\frac{1}{5} - \frac{s}{25} - \frac{s^2}{125}.$$

Then

$$\begin{aligned} h_2(s) &= (-2+s)h_1(s) + 3h_1(s) \mod s^3 \\ &= -\frac{1}{5} - \frac{6s}{25} - \frac{6s^2}{125}, \\ h_3(s) &= (-2+s)h_2(s) - 6h_1(s) \mod s^3 \\ &= \frac{8}{5} + \frac{13s}{25} - \frac{12s^2}{125}, \\ h_4(s) &= (-2+s)h_3(s) - 28h_1(s) \mod s^3 \\ &= \frac{12}{5} + \frac{42s}{25} + \frac{117s^2}{125}. \end{aligned}$$

As a check of computation, we verify that

$$(-2+s)h_4(s) - 24h_1(s) \mod s^3 = \frac{117s^3}{125} \mod s^3 = 0.$$

Thus it follows from $\begin{bmatrix} s^2 & s & 1 \end{bmatrix} W_1 = \begin{bmatrix} h_4 & h_3 & h_2 & h_1 \end{bmatrix}$ that

$$W_1 = \begin{bmatrix} \frac{117}{125} & -\frac{12}{125} & -\frac{6}{125} & -\frac{1}{125} \\ \frac{42}{25} & \frac{13}{25} & -\frac{6}{25} & -\frac{1}{25} \\ \frac{12}{5} & \frac{8}{5} & -\frac{1}{5} & -\frac{1}{5} \end{bmatrix}.$$

Similarly, for $\lambda_2 = 3$ we find that

$$\begin{aligned} h_1(s) &= \frac{1}{125}, \\ h_2(s) &= (3+s)h_1(s) + 3h_1(s) \mod s = \frac{6}{125}, \\ h_3(s) &= (3+s)h_2(s) - 6h_1(s) \mod s = \frac{12}{125}, \\ h_4(s) &= (3+s)h_3(s) - 28h_1(s) \mod s = \frac{8}{125}. \end{aligned}$$

and

$$(3+s)h_4(s) - 24h_1(s) \ \bmod\ s = 3 \cdot \frac{8}{125} - \frac{24}{125} = 0.$$

Then $\begin{bmatrix} 1 \end{bmatrix} W_2 = \begin{bmatrix} h_4 & h_3 & h_2 & h_1 \end{bmatrix}$ gives

$$W_2 = \begin{bmatrix} \frac{8}{125} & \frac{12}{125} & \frac{6}{125} & \frac{1}{125} \end{bmatrix}.$$

Finally, we have

$$V^{-1} = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} \frac{117}{125} & -\frac{12}{125} & -\frac{6}{125} & -\frac{1}{125} \\ \frac{42}{25} & \frac{13}{25} & -\frac{6}{25} & -\frac{1}{25} \\ \frac{12}{5} & \frac{8}{5} & -\frac{1}{5} & -\frac{1}{5} \\ \frac{8}{125} & \frac{12}{125} & \frac{6}{125} & \frac{1}{125} \end{bmatrix}.$$

Acknowledgment

This work was supported by the Research Committee of The Hong Kong Polytechnic University (Grant No. G-S744).

References

[1] D. K. FADDEEV AND V. N. FADDEEVA, *Computational Methods of Linear Algebra*, Freeman, San Francisco, 1963.

[2] H. K. GARG, *Digital Signal Processing Algorithms*, CRC Press, 1998.

[3] F. A. GRAYBILL, *Matrices with Applications to Statistics*, second ed., Wadsworth, Belmont, Calif., 1983.

[4] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, 1985.

[5] S. H. HOU, *On Leverrier-Faddeev Algorithm*, Proceedings of the Third Asian Technology Conference in Mathematics, Yang et al (Eds.), Springer Verlag, 399-403, 1998.

[6] T. KAILATH, *Linear Systems*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1980.

[7] D. KALMAN, *The Generalized Vandermonde Matrix*, Math. Mag., 57:15-21, 1984.

[8] A. Klinger, *The Vandermonde Matrix*, Amer. Math. Monthly, 74:571-574, 1967.

[9] N. Macon and A. Spitzbart, *Inverses of Vandermonde Matrices*, Amer. Math. Monthly, 65:95-100, 1958.

[10] C. Pozrikidis, *Numerical Computation in Science and Engineering*, Oxford University Press, 1998.