2007

On perfect conditioning of Vandermonde matrices on the unit circle

Lihu Berman
feuer@ee.technion.ac.il

Arie Feuer

Follow this and additional works at: http://repository.uwyo.edu/ela

Recommended Citation
DOI: https://doi.org/10.13001/1081-3810.1190

This Article is brought to you for free and open access by Wyoming Scholars Repository. It has been accepted for inclusion in Electronic Journal of Linear Algebra by an authorized editor of Wyoming Scholars Repository. For more information, please contact scholcom@uwyo.edu.
ON PERFECT CONDITIONING OF VANDERMONDE MATRICES ON THE UNIT CIRCLE∗

LIHU BERMAN† AND ARIE FEUER†

Abstract. Let $K, M \in \mathbb{N}$ with $K < M$, and define a square $K \times K$ Vandermonde matrix $A = A(\tau, \overrightarrow{n})$ with nodes on the unit circle: $A_{p,q} = \exp(-j2\pi pn_q\tau/K)$; $p, q = 0, 1, ..., K - 1$, where $n_q \in \{0, 1, ..., M - 1\}$ and $n_0 < n_1 < ... < n_{K-1}$. Such matrices arise in some types of interpolation problems. In this paper, necessary and sufficient conditions are presented on the vector $\overrightarrow{n}$ so that a value of $\tau \in \mathbb{R}$ can be found to achieve perfect conditioning of $A$. A simple test to check the condition is derived and the corresponding value of $\tau$ is found.

Key words. Vandermonde matrix, Condition number, Perfect conditioning.

AMS subject classifications. 15A12, 65F35.

1. Introduction and some notation. Vandermonde matrices with real nodes are often badly conditioned; see [8], [12] or [2]. However, this is not necessarily so when the nodes are complex. In fact, such matrices can even be unitary, hence perfectly-conditioned; e.g., the $N \times N$ DFT (Discrete Fourier Transform) matrix: $F_{a,b} = 1/\sqrt{N}\exp(-j2\pi ab/N)$.

Such matrices often arise in engineering problems such as interpolation, extrapolation, super-resolution and recovering of missing samples. One application which is of particular interest to us is interpolation from sub-Nyquist periodic non-uniform samples of band-limited multiband signals ([6], [10] and [4]). In all these applications the quality of the result depends heavily on the condition number of the aforementioned matrix (see more about this later).

This fact led several authors to investigate the conditioning of such matrices. For example, reference [5] studies the conditioning of Vandermonde matrices with nodes on the unit circle, given by the Van der Corput sequence. Reference [7] derives bounds on the condition number (through bounds on the singular values) of Vandermonde matrices with nodes on the unit circle, dependant on the minimum and maximum distances between the nodes of the generating row. Another reference([1]) considered the effect of increasing $N$ on the conditioning of $n \times N$ rectangular Vandermonde matrices with nodes in the unit disk.

Let us introduce some notation we use in the sequel. We denote by $S(f)$ the Fourier Transform of $s(t)$: $S(f) = \int_{\mathbb{R}} s(t) \exp(-j2\pi ft) dt$, where $j \triangleq \sqrt{-1}$. $A^H$ is the conjugate transpose of $A$ and $\kappa(A) \triangleq \|A\| \|A^{-1}\|$ denotes the condition number of $A$ and varies according to particular matrix norm chosen. Using the spectral norm leads to: $\kappa(A) \triangleq \sigma_{\max}(A) / \sigma_{\min}(A)$ where $\sigma_{\max}(A)$ (respectively: minimal) singular value of $A$.

∗Received by the editors 3 November 2006. Accepted for publication 8 July 2007. Handling Editor: Michael Neumann.
†Department of Electrical Engineering, Technion, Haifa, Israel (feuer@ee.technion.ac.il).
2. Motivation. Define the spectral support of \( s(t) \) as the set: \( \chi(s) = \{ f \in \mathbb{R} \mid S(f) \neq 0 \} \). A signal is band-limited if its spectral support is bounded: \( \chi(s) \subset (f_{\text{min}}, f_{\text{max}}) \). It is well known ([11]) that such a signal can be perfectly reconstructed from its uniform samples \( s(nT) \) taken at the Nyquist rate or higher: \( T^{-1} \geq (f_{\text{max}} - f_{\text{min}}) \).

A signal \( s(t) \) is called multi-band if its (bounded) spectral support is not connected; i.e. its spectral support consists of a union of a finite set of disjoint intervals: \( \chi(s) \subseteq \bigcup_{j=1}^{J} \left( f_{\text{min}} + \frac{\alpha_j}{M} (f_{\text{max}} - f_{\text{min}}), f_{\text{min}} + \frac{(\alpha_j+1)}{M} (f_{\text{max}} - f_{\text{min}}) \right) \).

In this case, it can be shown ([6] or [3]) that the signal \( s(t) \) can be perfectly reconstructed from its periodic non-uniform samples \( \bigcup_{p=0}^{K-1} s(nMT + t_p) \), where the average \( \frac{K}{M} T^{-1} \) is lower than the Nyquist rate.

A sampling equation for the aforementioned case is: \( \overrightarrow{G}(f) = A \overrightarrow{S}(f) \) where the \( K \)-dimensional vector \( \overrightarrow{G}(f) \) relates to the sampled signal, the \( K \)-dimensional vector \( \overrightarrow{S}(f) \) relates to the original signal and the \( K \times K \) matrix is given by: \( A_{p,q} = \exp(-j2\pi p t_p n_q / M); p, q = 0, 1, ..., K - 1 \). Note that the signal \( s(t) \) can be recovered if and only if the matrix \( A \) is invertible. The columns of \( A \) are determined by the spectral support, given by \( \overrightarrow{n} = [n_0, n_1, ..., n_{K-1}] \) and its rows are determined by the sampling pattern given by \( \overrightarrow{t} = [t_0, t_1, ..., t_{K-1}] \). The choice \( t_p = M/p \pi \) for the sampling pattern results in \( A \) being a Vandermonde matrix.

As a motivating remark we wish to point out that different sampling patterns lead to different quality of reconstruction in the presence of noise. Considering the set of linear equations above, \( \overrightarrow{G}(f) = A \overrightarrow{S}(f) \), it is well known that an error \( \overrightarrow{\Delta G}(f) \) in the data \( \overrightarrow{G}(f) \), will propagate to an error \( \overrightarrow{\Delta S}(f) \) in the solution \( \overrightarrow{S}(f) \). This error is bounded by: \( \frac{\overrightarrow{\Delta S}(f)}{\overrightarrow{S}(f)} \leq \kappa(A) \frac{\overrightarrow{\Delta G}(f)}{\overrightarrow{G}(f)} \), where \( \kappa(A) \) is the condition number of \( A \); see e.g., [13] for a more detailed discussion. This observation motivates our interest in the condition number of \( A \).

2.1. Problem Statement. Given \( K < M \in \mathbb{N} \) and \( \overrightarrow{n} = [n_0, n_1, ..., n_{K-1}] \), where \( n_q \in \{0,1,...,M-1\} ; n_0 < n_1 < ... < n_{K-1} \). Define the \( K \times K \) Vandermonde matrix \( A(\tau) \) as

\[
A_{p,q}(\tau) = e^{-j2\pi p n_q \tau / \kappa}; \quad 0 \leq p, q \leq K - 1.
\]

Find conditions on \( \overrightarrow{n} \) which guarantee the existence of \( \tau_{\text{opt}} \in \mathbb{R} \) so that \( \kappa(A(\tau_{\text{opt}})) = 1 \) (namely, for which the resulting matrix is perfectly conditioned) and find \( \tau_{\text{opt}} \).

3. Results. While the following result seems quite intuitive (as noted in [5]) we did not find anywhere in the literature a formal statement and proof for it. So, for the sake of completeness we decided to include both here.

**Lemma 3.1.** Let \( A(\tau) \) be as in (2.1). Then \( \kappa(A(\tau)) = 1 \) if and only if the elements of the generating row, \( \left\{ e^{-j2\pi p n_q / \kappa} \right\}_{q=0}^{K-1} \) are uniformly spread on the unit circle.
Perfect Conditioning of Vandermonde Matrices

Proof. Note first that by (2.1)

\[
(A^H (\tau) A (\tau))_{p,q} = \sum_{i=0}^{K-1} e^{\frac{2\pi i (n_p - n_q)}{K}} = \begin{cases}
K & \text{if } \tau (n_p - n_q) = mK; \ m \in \mathbb{Z} \\
1-e^{\frac{2\pi i (n_p - n_q)}{K}} & \text{otherwise.}
\end{cases}
\tag{3.1}
\]

Further note that for any square matrix \( A \)
\[
\kappa (A) = 1 \iff A^H A = cI \ \text{for some } c \in \mathbb{R}.
\tag{3.2}
\]

Combining (3.1) and (3.2) we get
\[
\kappa (A (\tau)) = 1 \iff \{ \tau (n_p - n_q) \in \mathbb{Z} \ \text{and} \ \tau (n_p - n_q) \not\equiv 0 \pmod{K} \}
\text{for all } 0 \leq p \neq q \leq K - 1
\tag{3.3}
\]
leading to
\[
\kappa (A (\tau)) = 1 \iff \tau n_p = \varepsilon + r_p K + s_p \ \text{where } \varepsilon \in \mathbb{R}; \ r_p, s_p \in \mathbb{Z}
\text{and } 0 \leq s_p \neq s_q \leq K - 1 \ \text{for all } 0 \leq p \neq q \leq K - 1,
\tag{3.4}
\]
which readily implies that \( \kappa (A (\tau)) = 1 \) if and only if the points \( \{ e^{-j \frac{2\pi n_p}{K}} \}^{K-1}_{q=0} \) are spread uniformly on the unit circle. \( \square \)

Clearly, the set of points \( \{ e^{-j \frac{2\pi n_p}{K}} \}^{K-1}_{q=0} \) is a rotated version of the set \( \{ e^{-j \frac{2\pi n_q}{K}} \}^{K-1}_{q=0} \). Hence, we can assume w.l.o.g. that \( n_0 = 0 \). Furthermore, we readily observe that the points \( \{ e^{-j \frac{2\pi n_q}{K}} \}^{K-1}_{q=0} \) are spread uniformly if and only if \( \{ \tau n_q \}^{K-1}_{q=0} \) is a complete residue system \(^1\) mod \( K \).

We can now establish our main result.

**Theorem 3.2.** Let \( A (\tau) \) be as in (2.1) and let \( Q = \gcd \left( \left\{ n_q \right\}^{K-1}_{q=1} \right) \). Then, there exists \( \tau \in \mathbb{R} \) such that \( \kappa (A (\tau)) = 1 \) if and only if \( \left\{ \left\{ \frac{n_q}{Q} \right\}^{K-1}_{q=1} \right\}^{K-1}_{q=1} \) is a complete residue system mod \( K \).

**Proof.** We know already that \( \kappa (A (\tau)) = 1 \) if and only if \( \{ \tau n_q \}^{K-1}_{q=0} \) is a complete residue system mod \( K \). Thus, it suffices to show that there exists a \( \tau \in \mathbb{R} \) such that \( \{ \tau n_q \}^{K-1}_{q=0} \) is a complete residue system mod \( K \) if and only if \( \left\{ \left\{ \frac{n_q}{Q} \right\}^{K-1}_{q=1} \right\}^{K-1}_{q=1} \) is a complete residue system mod \( K \).

\(^1\) \( \{ r_q \}^{K-1}_{q=0} \) is called a complete residue system mod \( K \) if for any \( a \in \mathbb{Z} \) there exists a unique \( r_q \) such that \( a \equiv r_q \pmod{K} \). meaning that \( r_p \equiv r_q \pmod{K} \iff p = q \) for all \( 0 \leq p, q \leq K - 1 \).
Sufficiency is straightforward as, given that \( \left\{ \left( \frac{n_q}{Q} \right) \right\}_{q=1}^{K-1} \) is a complete residue system mod \( K \), choose \( \tau = \frac{1}{Q} \) and we are done.

For necessity, assume there exists \( \tau \in \mathbb{R} \) such that \( \left\{ \tau n_q \right\}_{q=0}^{K-1} \) is a complete residue system mod \( K \). Then we have

\[
\tau n_q = \tau Q \frac{n_q}{Q} \in \mathbb{Z} \Rightarrow \tau Q = R \in \mathbb{Z}
\]

\[
\Rightarrow \tau = \frac{R}{Q}.
\] (3.5)

This means, by assumption, that \( \left\{ \left( \frac{n_q}{Q} \right) \right\}_{q=1}^{K-1} \) is a complete residue system mod \( K \).

It is known (see e.g., Theorem 2.1.2 in [9]) that for any \( a, b, c \in \mathbb{Z} \) and \( n \in \mathbb{N} \)

\[
ac \equiv bc \pmod{n} \Leftrightarrow a \equiv b \pmod{\frac{n}{g}} \text{ where } g = \gcd(c, n) \quad (3.6)
\]

So, taking \( a = \frac{n_p}{Q}, b = \frac{n_q}{Q}, c = R \) and \( n = K \) and applying (3.6) we get

\[
R \frac{n_p}{Q} \equiv R \frac{n_q}{Q} \pmod{K} \Leftrightarrow \frac{n_p}{Q} \equiv \frac{n_q}{Q} \pmod{\frac{K}{g}} \text{ where } g = \gcd(R, K) \quad (3.7)
\]

Then, as a complete residue system mod \( \frac{K}{g} \) contains exactly \( \frac{K}{g} \) distinct elements, if \( \frac{K}{g} < K \) there must be at least two integers \( 0 \leq p \neq q \leq K - 1 \) such that \( \frac{n_p}{Q} \equiv \frac{n_q}{Q} \pmod{\frac{K}{g}} \). However, by (3.7) this implies that \( R \frac{n_p}{Q} \equiv R \frac{n_q}{Q} \pmod{K} \) for some \( p \neq q \) and contradicts the fact that \( \left\{ \left( \frac{Kn_q}{Q} \right) \right\}_{q=1}^{K-1} \) is a complete residue system mod \( K \). This necessarily means that \( g = 1 \) and by (3.7), that \( \left\{ \left( \frac{n_q}{Q} \right) \right\}_{q=1}^{K-1} \) is a complete residue system mod \( K \), which completes the proof of the theorem. \( \square \)

The above theorem also suggests a simple test for the existence and calculation of a \( \tau_{\text{opt}} \in \mathbb{R} \) such that \( \kappa(A(\tau_{\text{opt}})) = 1 \). Assuming the condition in Theorem 3.2 is satisfied we can choose \( \tau_{\text{opt}} = \frac{1}{Q} \) (note that this choice is not unique since \( -\tau_{\text{opt}} \) and \( \tau_{\text{opt}} + nK \) will also give the same condition number).

4. Conclusion. This paper presents necessary and sufficient conditions on the vector \( \vec{n} \) such that there exists \( \tau \in \mathbb{R} \) for which the matrix \( A_{p,q}(\tau) = e^{-j \frac{2\pi p n q}{K}}/K \), \( p, q = 0, 1, ..., K - 1 \) is perfectly conditioned.
REFERENCES


