1 Introduction

In numerical analysis, it is usually important to know the stability of different methods. To this end, the idea of the condition number is introduced to measure how much the output can change for a small change in the input argument. A problem with a high condition number is said to be ill-conditioned, where the output is sensitive to changes or errors in the input. Vandermonde matrices can evaluate a polynomial at a set of points, and thus are widely used in solving interpolation problems. However, they are typically very ill-conditioned. This report aims to explore the reason behind their ill-conditioning.

2 Vandermonde Matrices

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. While they are indexed from 1 in the above definition, sometimes they are indexed from 0.

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 $(x_1, y_1), ..., (x_n, y_n)$. It is equivalent to solving the system $V^T a = f$.

The inverse of a Vandermonde matrix is useful in computing $\kappa(V) = ||V^{-1}||||V||$ (shown later). To find the inverse, 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

$$\sum_{j=1}^{n} w_{ij} x_k^{j-1} = \delta_{ik}, \quad k = 1, 2, ..., n,$$

where δ_{ij} is the Kronecker delta (equal to 1 if i=j and 0 otherwise).

We see that this polynomial is the Lagrange basis polynomial:

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

We deduce that

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

where $\sigma_k(y_1, \dots, y_n)$ denotes the sum of all distinct products of k of the arguments $y_1, \dots, y_n$.

3 Condition Number

A condition number of a problem measures the sensitivity of the solution to small perturbations in the input data. The condition number depends on the problem and the input data, on the norm used to measure size, and on whether perturbations are measured in an absolute or a relative sense.

The condition number of a matrix is defined to be

$$\kappa(A) = ||A^{-1}||||A||.$$

It satisfies that $\kappa(A) = ||A^{-1}||||A|| \geq ||A^{-1}A|| = 1$.

If $||\cdot||$ is the matrix norm induced by the Euclidean norm (L^2 norm), the condition number of a matrix can be written as

$$\kappa(A) = \frac{\sigma_{max}(A)}{\sigma_{min}(A)}$$

where $\sigma_{max}(A)$ and $\sigma_{min}(A)$ are the maximal and minimal singular values of A respectively.

The following examples illustrate the condition numbers for some simple matrices.

$$A = I \;\; => \;\; \kappa(A) = \infty$$

$$A = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \;\; => \;\; \kappa(A) = \infty$$

$$A = \begin{bmatrix} I & \cdot \\ \cdot & 10^{-k} \end{bmatrix} => \;\; \kappa(A) = 10^k$$

The range of condition number is $[1, \infty)$. A problem with a low condition number (i.e. $\kappa(A)$ is close to 1) is said to be well-conditioned, while a problem with a high condition number is said to be ill-conditioned.

4 Why Vandermonde Matrices are Ill-conditioned

For a non-singular Vandermonde matrix V, denote that

$$l_\dagger = \max_{1 \le i \le n-1} |l_i|, \quad |V| = \max\{1, l_\dagger^{n-1}\}.$$

Then a Vandermonde matrix $V = (l_i^j)_{i,j=0}^{n-1}$, $||V||$ can be bounded by

$$(i)|V| \le ||V|| \le n|V|, \quad (ii)||V|| \ge \sqrt{n}.$$

They are proved in the book Matrix Computations by Johns Hopkins University Press.

It also holds that

$$(i)\sigma_n(V) \le \sqrt{n},$$

$$(ii)\sigma_n(V) \le \nu^{(1-k)}\sqrt{k}\max\{k, \frac{\nu}{\nu-1}\} \text{ if } \frac{1}{|L_i|} \ge \nu > 1 \text{ for } i = 0, 1, ..., k-1.$$

Those are following the theorem that: $\sigma_j(M)$ is equal to the distance $||M - M_{j1}||$ between M and its closest approximation by a matrix M_{j-1} of rank at most $j-1$.

The collieries are that for a Vandermonde matrix $V = (l_i^j)_{i,j=0}^{n-1}$,

$$\kappa(V) \ge \max\{1, \frac{l_\dagger^{n-1}}{\sqrt{n}}\}$$

$$\kappa(V) \ge \frac{|V|\nu^{k-1}}{\sqrt{k}\max\{k, \frac{\nu}{\nu-1}\}} \text{ if } \frac{1}{|l_i|} \ge \nu > 1 \text{ for } i = 0, 1, ..., k-1$$

We can see from the first inequality that $\kappa(V)$ is exponential in number of all nodes n if $l_\dagger \ge \nu > 1$.
We can see from the second inequality that $\kappa(V)$ is exponential in k if $\frac{1}{|l_i|} \ge \nu > 1$ for at least k nodes.

Another paper considers the "L^n-norm" form of the condition number and gives some common examples illustrating the exponential growth.

Let V_n be the Vandermonde matrix obtained from Lagrange interpolation of degree n.

Case 1: For Harmonic nodes $X_i = 1/i, i = 1, 2, ..., n$,

$$\kappa(V_n) > n^{(n+1)}.$$

The condition number grows exponentially.

Case 2: For Equidistant nodes on $[0, 1]$: $x_i = \frac{i-1}{n-1}, i = 1, 2, ..., n$,

$$\kappa(V_n) \approx \frac{\sqrt{2}}{4\pi} 8^n.$$

The condition number also grows exponentially, but the growth rate is 8, which is much smaller among them.

Case 3: For Equidistant nodes on $[-1, 1]$: $X_i = 1 - \frac{2(i-1)}{n-1}, i = 1, 2, ..., n$,

$$\kappa(V_n) \approx \frac{1}{\pi} e^{-\frac{1}{4}\pi} e^{n(\frac{1}{4}\pi + \frac{1}{2}\ln 2)}.$$

The condition number also grows exponentially, but the growth rate is $e^{n(\frac{1}{4}\pi+\frac{1}{2}\ln 2)} \approx 3.1017$.

Case 4: For Chebyshev nodes $X_i = \cos(\frac{(2i-1)\pi}{2n}), i = 1, 2, ..., n$,

$$\kappa(V_n) \approx \frac{3^{\frac{3}{4}}}{4} (1 + \sqrt{2})^n.$$

The condition number also grows exponentially, but the growth rate is $1 + \sqrt{2} \approx 2.4142$, which is the smallest.

The four cases show the decreasing severity of ill-conditioning. We can see that for all choices of interpolation nodes, the condition numbers exhibit exponential growth, but vary greatly in the growth rate.

5 Discussion

The section above shows the idea to prove that the condition numbers for Vandermonde matrices have exponential growth for all real sets of nodes. Since the number of nodes n is moderately large in most cases, we conclude that those Vandermonde matrices are very poorly conditioned.

The ill-conditioning has a narrow class of notable exceptions when the nodes are not real, such as the matrice of the discrete Fourier transform (DFT) below.

$$W = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \cdots & \omega^{2(N-1)} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \cdots & \omega^{3(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \omega^{3(N-1)} & \cdots & \omega^{(N-1)(N-1)} \end{bmatrix}$$

where $\omega = e^{-2\pi i/N}$ is a primitive Nth root of unity.

The matrices of DFT are obviously Vandermonde matrices. They are defined by a cyclic sequence of nodes, equally spaced on a circle. In fact, the matrices of DFT are perfectly conditioned.