Accurate Inverses Through Accurate Minors for Structured Matrices

James Demmel, Plamen Koev

University of California - Berkeley

- **Cauchy matrix:**

$$C(i,j) = \frac{1}{x_i + y_j}$$

- **Vandermonde matrix:**

$$V = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ & & \ddots & \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{bmatrix} \quad x_1 > x_2 > \dots > x_n > 0 \quad \textbf{TP}$$

- **Generalized Vandermonde matrix:**

$$G = \begin{bmatrix} x_1^{a_1} & x_2^{a_1} & \dots & x_n^{a_1} \\ x_1^{a_2} & x_2^{a_2} & \dots & x_n^{a_2} \\ & & \ddots & \\ x_1^{a_n} & x_2^{a_n} & \dots & x_n^{a_n} \end{bmatrix} \quad \text{and} \quad x_1 > \dots > x_n > 0; a_i \in Z$$

- **GOAL:**

  Computing C^{-1}, V^{-1} and G^{-1} to High Relative Accuracy

- **HISTORY:**
  - **Bjorck-Pereyra 1970: Accurate V^{-1} in $O(n^3)$ time.**
  - **Olshevsky 1995: Extends this to Cauchy**
  - **Demmel 1997: Accurate SVD for Cauchy, Vandermonde etc. in $O(n^3)$ time**

INSPIRATION

Necessary and Sufficient Conditions – Demmel:

- **Being able to compute $|\det(A)| = \sigma_1\sigma_2\ldots\sigma_n$ accurately is a necessary condition for accurate SVD.**
- **Similar sufficient condition for accurate A^{-1}: Enough to compute n^2+1 minors (Cramer's Rule)**
- **Sufficient condition for accurate SVD: Enough to compute $O(n^3)$ minors (Demmel 97)**

CONTRIBUTIONS:

- **By using Cramer's Rule we obtain new $O(n^3)$ algorithm for accurate inversion of Cauchy, Vandermonde and some GENERALIZED Vandermonde matrices**
- **For those matrices being able to compute n^2+1 minors is also a NECESSARY for computing an accurate inverse in $O(n^3)$ time**
- **We match the current accuracy and speed benchmark for inverting Vandermondes.**
- **We make some progress in computing accurate SVDs of Generalized Vandermonde matrices**

Cauchy Matrices

$$C(i,j) = \frac{1}{x_i + y_j}$$

$$(C^{-1})_{ij} = \frac{\Pi_{k=1}^{n}(x_j + y_k)(x_k + y_i)}{(x_i + y_j)\,\Pi_{k=1,k\neq j}^{n}(x_j - x_k)\,\Pi_{k=1,k\neq j}^{n}(y_i - y_k)}$$

C^{-1} Computable accurately in $O(n^3)$ time

Vandermonde Matrices

$$V = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ & & \ddots & \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{bmatrix} \qquad x_1 > x_2 > \ldots > x_n > 0$$

THE MINORS:

$$G = \begin{bmatrix} x_1^{a_1} & x_2^{a_1} & \dots & x_n^{a_1} \\ x_1^{a_2} & x_2^{a_2} & \dots & x_n^{a_2} \\ & & \ddots & \\ x_1^{a_n} & x_2^{a_n} & \dots & x_n^{a_n} \end{bmatrix} \qquad x_1 > \ldots > x_n > 0; \quad a_i \in Z$$

$$\det(G) = \det(V) \cdot s_\lambda(x_1, x_2, ..., x_n),$$

- $\det(V) = \Pi_{i>j}(x_i - x_j)$ **- computable accurately**
- $s_\lambda(x_1, x_2, ..., x_n)$ **- is called SCHUR FUNCTION**

The Schur Function

- **is a polynomial with positive integer coefficients and is therefore computable accurately**
- **only depends on the PARTITION:**

$$\lambda = (a_n - (n-1), a_{n-1} - (n-2), ..., a_2 - 1, a_1 - 0)$$

- **has exponentially many terms -** $O(n^{|\lambda|})$
(for a partition $\lambda = (\lambda_1, ..., \lambda_n); \;\; \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$ **we have** $|\lambda| := \lambda_1 + ... + \lambda_n$**)**
- **the** $(n-1) \times (n-1)$ **minors of Vandermonde correspond to partitions:** $\lambda_m = (1, 1, ..., 1); \; |\lambda_m| = m$
- **Recursive formula (McDonald):**

$$s_\lambda(x_1, ..., x_n, y_1, ..., y_m) = \sum_{\mu < \lambda} s_{\lambda/\mu}(x_1, ..., x_n) \cdot s_\mu(y_1, ..., y_m)$$

- **Let** $s_{mk} := s_{\lambda_m}(x_1, ..., x_k)$**,**
McDonald $\Rightarrow \;\; s_{mk} = s_{m,k-1} + s_{m-1,x-1} \cdot x_n$

$$\begin{array}{ccccccccc}
s_{01} & & s_{11} & & \mathbf{0} & & \ldots & & \mathbf{0} \\
 & & \downarrow & & & & & & \\
s_{02} & \rightarrow & s_{12} & \rightarrow & s_{22} & \rightarrow & & & \mathbf{0} \\
 & & \downarrow & & \downarrow & & & & \\
s_{03} & \rightarrow & s_{13} & \rightarrow & s_{23} & \rightarrow & & \rightarrow & \mathbf{0} \\
 & & \downarrow & & \downarrow & & & & \\
 & & \vdots & & \vdots & & \ddots & & \\
 & & \downarrow & & \downarrow & & & & \downarrow \\
s_{0,n-1} & \rightarrow & s_{1,n-1} & \rightarrow & s_{2,n-1} & \rightarrow & \cdots & \rightarrow & s_{n-1,n-1}
\end{array}$$

- **(For** $(V^{-1})_{ij}$ **we need** $s_{i,n-1}$**, i.e. the last row)**
Total running time $O(n^3)$ **for accurate** V^{-1}

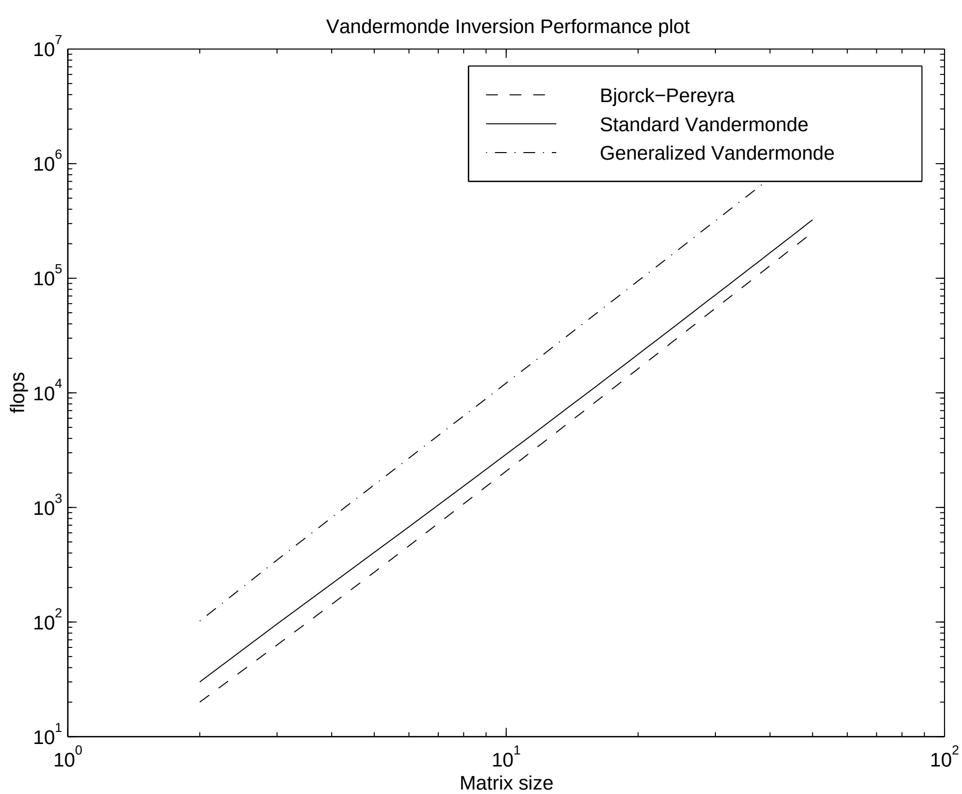
Inverting Some Generalized Vandermondes

$$G = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ & & \ddots & \\ x_1^{n-2} & x_2^{n-2} & \dots & x_n^{n-2} \\ x_1^n & x_2^n & \dots & x_n^n \end{bmatrix} \quad x_1 > x_2 > \ldots > x_n > 0$$

- **The partitions corresponding to $(n-1) \times (n-1)$ minors are $\lambda'_m = (2, \underbrace{1, 1, ..., 1}_{m-1})$**
- **Let $g_m(x_1, ..., x_n) := s_{\lambda'}(x_1, ..., x_n)$**
  McDonald: Similar recursive relationship

$$g_m(x_1, ..., x_n) = g_{m-1}(x_1, ..., x_{n-1}) + s_{m-1,n-1}(s_{1,n-1} + x_n)x_n$$

- **Again, invertible accurately in $O(n^3)$ time**



Determinant of any Generalized Vandermonde

- **Necessary condition for an accurate SVD, still very hard**
- **All minors are also Generalized Vandermondes**
- **Doable in principle independent of** $cond(G)$ **(but in exponential amount of time)**
- $\det(G) = \det(V) \cdot s_\lambda(x_1, ..., x_n)$
- s_λ **has** $O(n^{|\lambda|})$ **terms**
- **Use divide-and-conquer with**

$$s_\lambda(x_1, ..., x_n, y_1, ..., y_m) = \sum_{\mu<\lambda} s_{\lambda/\mu}(x_1, ..., x_m) \cdot s_\mu(y_1, ..., y_m)$$

- **New Result: With the recursive formula: New bound on cost:**

$$O(n^{\log \lambda_1 + ... + \log \lambda_n})$$

- **Still exponential, but exponentially better than**

$$O(n^{|\lambda|}) = O(n^{\lambda_1 + ... + \lambda_n})$$

- `http://www.math.berkeley.edu/~plamen`