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:
  \[
  V = \begin{bmatrix}
  1 & 1 & \ldots & 1 \\
  x_1 & x_2 & \ldots & x_n \\
  x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1}
  \end{bmatrix}
  \quad \text{if } x_1 > x_2 > \ldots > x_n > 0 \quad \text{TP}
  \]

- Generalized Vandermonde matrix:
  \[
  G = \begin{bmatrix}
  x_1^{a_1} & x_2^{a_1} & \ldots & x_n^{a_1} \\
  x_1^{a_2} & x_2^{a_2} & \ldots & x_n^{a_2} \\
  \vdots & \vdots & \ddots & \vdots \\
  x_1^{a_n} & x_2^{a_n} & \ldots & x_n^{a_n}
  \end{bmatrix}
  \quad \text{and} \quad x_1 > \ldots > x_n > 0; a_i \in \mathbb{Z}
  \]

- 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

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

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

Vandermonde Matrices

\[
V = \begin{bmatrix}
1 & 1 & \ldots & 1 \\
x_1 & x_2 & \cdots & x_n \\
\vdots & & \ddots & \vdots \\
x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1}
\end{bmatrix}
\]
\( x_1 > x_2 > \ldots > x_n > 0 \)

THE MINORS:

\[
G = \begin{bmatrix}
x_1^{a_1} & x_2^{a_1} & \ldots & x_n^{a_1} \\
x_1^{a_2} & x_2^{a_2} & \ldots & x_n^{a_2} \\
\vdots & \vdots & \ddots & \vdots \\
x_1^{a_n} & x_2^{a_n} & \cdots & x_n^{a_n}
\end{bmatrix}
\]
\( x_1 > \ldots > x_n > 0; \ a_i \in \mathbb{Z} \)

\[
\det(G) = \det(V) \cdot s_\lambda(x_1, x_2, \ldots, x_n),
\]

- \( \det(V) = \prod_{i>j}(x_i - x_j) \) - computable accurately
- \( s_\lambda(x_1, x_2, \ldots, x_n) \) - is called SCHUR FUNCTION
The Schur Function

- is a polynomial with positive integer coefficients and is therefore computable accurately

- only depends on the PARTITION:

\[ \lambda = (a_n - (n - 1), a_{n-1} - (n - 2), \ldots, a_2 - 1, a_1 - 0) \]

- has exponentially many terms - \( O(n^{|\lambda|}) \)
  (for a partition \( \lambda = (\lambda_1, \ldots, \lambda_n); \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0 \) we have \( |\lambda| := \lambda_1 + \ldots + \lambda_n \))

- the \((n - 1) \times (n - 1)\) minors of Vandermonde correspond to partitions: \( \lambda_m = (1, 1, \ldots, 1); |\lambda_m| = m \)

- Recursive formula (McDonald):

\[ s_{\lambda}(x_1, \ldots, x_n, y_1, \ldots, y_m) = \sum_{\mu \leq \lambda} s_{\lambda/\mu}(x_1, \ldots, x_n) \cdot s_{\mu}(y_1, \ldots, y_m) \]

- Let \( s_{mk} := s_{\lambda_m}(x_1, \ldots, x_k) \),
  McDonald \( \Rightarrow \)

\[
\begin{align*}
  s_{01} & \rightarrow s_{11} & 0 & \ldots & 0 \\
  \downarrow & & & & \\
  s_{02} & \rightarrow s_{12} & \rightarrow s_{22} & \rightarrow & 0 \\
  \downarrow & & & & \\
  s_{03} & \rightarrow s_{13} & \rightarrow s_{23} & \rightarrow & 0 \\
  \downarrow & & & & \\
  \vdots & \vdots & \ddots & \vdots & \vdots \\
  \downarrow & & & & \\
  s_{0,n-1} & \rightarrow s_{1,n-1} & \rightarrow s_{2,n-1} & \rightarrow \ldots & \rightarrow s_{n-1,n-1}
\end{align*}
\]

- (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

\[
G = \begin{bmatrix}
1 & 1 & \ldots & 1 \\
x_1 & x_2 & \ldots & x_n \\
x_1^{n-2} & x_2^{n-2} & \ldots & x_n^{n-2} \\
x_1^n & x_2^n & \ldots & x_n^n
\end{bmatrix}
\quad \text{with } x_1 > x_2 > \ldots > x_n > 0
\]

- The partitions corresponding to \((n - 1) \times (n - 1)\) minors are \(\lambda_m' = (2, 1, 1, \ldots, 1)_{m-1}\)

- Let \(g_m(x_1, \ldots, x_n) := s_{\lambda'}(x_1, \ldots, x_n)\)

McDonald: Similar recursive relationship

\[
g_m(x_1, \ldots, x_n) = g_{m-1}(x_1, \ldots, x_{n-1}) + s_{m-1,n-1}(s_{1,n-1} + x_n)x_n
\]

- 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 \( \text{cond}(G) \) (but in exponential amount of time)
- \( \det(G) = \det(V) \cdot s_\lambda(x_1, \ldots, x_n) \)
- \( s_\lambda \) has \( O(n^{|\lambda|}) \) terms
- Use divide-and-conquer with
  \[
  s_\lambda(x_1, \ldots, x_n, y_1, \ldots, y_m) = \sum_{\mu < \lambda} s_{\lambda/\mu}(x_1, \ldots, x_m) \cdot s_\mu(y_1, \ldots, y_m)
  \]
- New Result: With the recursive formula: New bound on cost:
  \[
  O\left( n^{\log \lambda_1 + \ldots + \log \lambda_n} \right)
  \]
- Still exponential, but exponentially better than
  \[
  O(n^{|\lambda|}) = O(n^{\lambda_1 + \ldots + \lambda_n})
  \]
- http://www.math.berkeley.edu/~plamen