Accurate and Efficient Matrix Computations with Generalized Vandermonde Matrices using Schur Functions

---

Plamen Koev
Department of Mathematics
UC Berkeley

Joint work with James Demmel

Supported by NSF and DOE

Vandermonde Matrices

- **TP Vandermonde Matrix:**

$$V = \begin{bmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ 1 & x_2 & \dots & x_2^{n-1} \\ & & \ddots & \\ 1 & x_n & \dots & x_n^{n-1} \end{bmatrix},$$

  where $x_1 > x_2 > \cdots > x_n > 0$**.**

- **TP Generalized Vandermonde Matrix:**

$$G_\lambda = \begin{bmatrix} x_1^{\lambda_1} & x_1^{1+\lambda_2} & \dots & x_1^{n-1+\lambda_n} \\ x_2^{\lambda_1} & x_2^{1+\lambda_2} & \dots & x_2^{n-1+\lambda_n} \\ & & \ddots & \\ x_n^{\lambda_1} & x_n^{1+\lambda_2} & \dots & x_n^{n-1+\lambda_n} \end{bmatrix},$$

  where

  - $x_1 > x_2 > \cdots > x_n > 0$
  - $\lambda_n \geq \lambda_{n-1} \geq \cdots \geq \lambda_0$

- **Notoriously illconditioned**

GOALS

Given a matrix A, compute to high relative accuracy in poly time:

- $\det(A)$
  - **not trivial even for tridiagonals**
- A^{-1}
- **Solve** $Ax = b$
- **LU from GENP, GEPP, GECP**
- **SVD**

Model of Arithmetic: $1+\delta$

- $\mathrm{fl}(a \otimes b) = (a \otimes b)(1+\delta)$
- **Unbounded Exponents**
- **Expressions we can compute without losing relative accuracy:**
  - **Products and quotients**
  - **Sums of positive quantities**
  - $x_i \pm x_j$**, where** x_i **and** x_j **are initial data**
- **Proof:** $1+\delta$ **factors can be factored out**
- **We will compute everything using allowable expressions**

Statement of Results

- **Can compute G^{-1} accurately for certain generalized Vandermondes in $O(n^3)$ time**
- $\det(G)$ **in $O(n^2)$ time**
- **LU of G from GENP, GEPP in $O(n^3)$ time, but who cares...**
- **Can solve $Gx = b$ with "Björck-Pereyra accuracy", but speed $O(n^3)$, not $O(n^2)$**

Type of Matrix	$\det(A)$	A^{-1}	**Any minor**	**GENP / GENP**	**GECP**	**SVD**
Cauchy	$O(n^2)$	$O(n^2)$	$O(n^2)$	$O(n^3)$	$O(n^3)$	$O(n^3)$
Vandermonde	$O(n^2)$	**NONE**	**NONE**	**NONE**	**NONE**	$O(n^3)$
Totally Positive (TP) Vandermonde	$O(n^2)$	$O(n^3)$	**EXP**	$O(n^3)$	**EXP**	$O(n^3)$
TP Generalized Vandermonde, $\lvert\lambda\rvert = \lambda_1 + O(1)$	$O(\lambda_1 n^2)$	$O(\lambda_1 n^3)$	**EXP**	$O(\lambda_1 n^3)$	**EXP**	**EXP**
TP Generalized Vandermonde, any λ	$f(n, \lambda)$	**EXP**	**EXP**	**EXP**	**EXP**	**EXP**

$$f(n,\lambda) = O(n^{1+\log\lambda_1+\log\lambda_2+\cdots+\log\lambda_n})\ \textbf{vs.}\ O(n^{|\lambda|}) = O(n^{\lambda_1+\lambda_2+\cdots+\lambda_n})$$

(Symmetrica)

Known Resuts - Björck-Pereyra Methods

$$V^{-1} = \begin{bmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ 1 & x_2 & \dots & x_2^{n-1} \\ & & \ddots & \\ 1 & x_n & \dots & x_n^{n-1} \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -x_1 & & \\ & 1 & -x_1 & \\ & & 1 & -x_1 \\ & & & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & & & \\ & 1 & -x_2 & \\ & & 1 & -x_2 \\ & & & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & -x_{n-1} \\ & & & 1 \end{bmatrix}$$

$$\cdot \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & -\frac{1}{x_n-x_1} & \frac{1}{x_n-x_1} \end{bmatrix} \cdots \begin{bmatrix} 1 & & & \\ -\frac{1}{x_2-x_1} & \frac{1}{x_2-x_1} & & \\ & -\frac{1}{x_3-x_2} & \frac{1}{x_3-x_2} & \\ & & -\frac{1}{x_n-x_{n-1}} & \frac{1}{x_n-x_{n-1}} \end{bmatrix}$$

- **Solving $Vz = b$ extremely accurate if b has "checkerboard pattern"**
- **Sign pattern result of Total Positivity, so only question is accuracy**
- **For SOLVING we could assemble V^{-1} and then form $V^{-1}b$ and the accuracy will be the same, but not speed**
- **Sign pattern of V^{-1} also alternating**
- **We will use Cramer's rule to compute inverses of generalized Vandermondes G, to which BP does not apply**

- **Then we can solve with BP accuracy by forming** $G^{-1}b$
- **Interesting: Even for** V**: Cramer's rule for** V^{-1}**:** $5/12n^3$**, BP:** $2n^3$

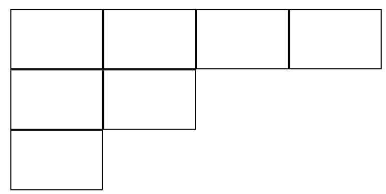
TP Vandermonde and Generalized Vandermonde Matrices

$$V = \begin{bmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ 1 & x_2 & \dots & x_2^{n-1} \\ & & \ddots & \\ 1 & x_n & \dots & x_n^{n-1} \end{bmatrix}, \qquad G_\lambda = \begin{bmatrix} x_1^{\lambda_1} & x_1^{1+\lambda_2} & \dots & x_1^{n-1+\lambda_n} \\ x_2^{\lambda_1} & x_2^{1+\lambda_2} & \dots & x_2^{n-1+\lambda_n} \\ & & \ddots & \\ x_n^{\lambda_1} & x_n^{1+\lambda_2} & \dots & x_n^{n-1+\lambda_n} \end{bmatrix},$$

where $x_1 > x_2 > \cdots > x_n > 0$**,** $\lambda_n \geq \lambda_{n-1} \geq \cdots \geq \lambda_0$

- **Def:** $\lambda = (\lambda_n, \lambda_{n-1}, ..., \lambda_0)$ **is partition of** $|\lambda| = \lambda_n + \cdots + \lambda_0$
- **Def: Young Diagrams** $\equiv$ **partitions:**

$$\lambda = (4, 2, 1) =$$



- **Def: Schur Function** $s_\lambda(x_1, \dots, x_n) = \det(G_\lambda)/\det(V)$**.**
- **Thm:** s_λ **is a polynomial with positive integer coefficients depending only on** λ **(MacDonald)**

Facts about TP Generalized Vandermonde Matrices

- **Recall:** $\det(G) = \det(V) \cdot s_\lambda(x_1, ..., x_n)$
- **Example:**

$$\det\left(\begin{bmatrix} 1 & x_1^2 & x_1^4 \\ 1 & x_2^2 & x_2^4 \\ 1 & x_3^2 & x_3^4 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}\right) \cdot (2x_1x_2x_3 + x_1^2x_2 + x_1x_2^2 + x_1^2x_3 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2)$$

- $\det(V) = \Pi_{i>j}(x_i - x_j)$ **is computable accurately and efficiently**
- s_λ **computable accurately** $(x_i > 0)$**, question is cost.**
- **Theorem:**

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

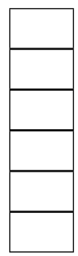
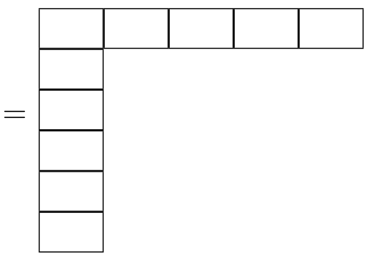
Allows recursive computation and Divide-and-Conquer approach.

- **Example:**

$$s_{(1,1)}(x_1, ..., x_n) = \sum_{i<j} x_i x_j = (x_1 + ... + x_n)x_2 + (x_2 + ... + x_n)x_3 + ... + x_{n-1}x_n$$

cost: $O(n)$**, although** $s_{(1,1)}$ **has** $O(n^2)$ **terms.**

- **Some** s_λ**'s are computable accurately and efficiently:**

$\lambda = (1, 1, 1, ..., 1) =$  **and** $\lambda = (m, 1, 1, ..., 1) =$ 

- **For Cramer's rule on a Vandermonde Matrix we need** $(n-1) \times (n-1)$ **minors:**

$$W_{ij} = \begin{bmatrix} 1 & x_1 & \dots & x_1^{j-2} & x_1^j & \dots & x_1^{n-1} \\ 1 & x_2 & \dots & x_2^{j-2} & x_2^j & \dots & x_2^{n-1} \\ & & \dots & & & \dots & \\ 1 & x_{i-1} & \dots & x_{i-1}^{j-2} & x_{i-1}^j & \dots & x_{i-1}^{n-1} \\ 1 & x_{i+1} & \dots & x_{i+1}^{j-2} & x_{i+1}^j & \dots & x_{i+1}^{n-1} \\ & & \dots & & & \dots & \\ 1 & x_n & \dots & x_n^{j-2} & x_n^j & \dots & x_n^{n-1} \end{bmatrix}$$

- **The partition that corresponds to W_{ij} is**

  $\lambda = (n-1-(n-2), n-2-(n-3), \ldots, j-(j-1), j-2-(j-2), \ldots, 1-1, 0) = (1, 1, ..., 1) = (1^{n-j})$.

- **The same trick applies to some generalized Vandermondes, e.g.**

$$G = \begin{bmatrix} 1 & x_1 & x_1^2 & & x_1^{n-2} & x_1^{n-1+m} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-2} & x_2^{n-1+m} \\ & & & \ddots & & \\ 1 & x_n & x_n^2 & & x_n^{n-2} & x_n^{n-1+m} \end{bmatrix}$$

- **Partition of G is $\lambda = (m)$**
- **Partitions of $(n-1) \times (n-1)$ minors $\lambda = (m, 1, 1, 1, ..., 1)$**
- **Cost of inversion: $O(mn^3)$**
- **Extends to $\lambda = (\lambda_1, \lambda_2, ...)$ as long as $\lambda_2 + \lambda_3 + ...$ is small ($O(1)$).**
- **G^{-1} will have "checkerboard" sign pattern**
- **Finally, with accurate inverses we can solve accurately**

Conclusions

- **Can compute accurate determinants and inverses of some Generalized Vandermonde Matrices in $O(n^3)$ time**
- **Can solve linear systems with these systems to high accuracy**

Open Problems

- **Inverting Any Generalized Vandermonde**
- **Accurate bidiagonal decomposition of G^{-1} like BP?**
- **Totally Positive Matrices in General**
- **Subtraction-free complexity of evaluating the Schur Function**