Matrices With Displacement Structure

James Demmel
Department of Mathematics
Department of Electrical Engineering and Computer Science
University of California - Berkeley
demmel@cs.berkeley.edu

Plamen Koev
Department of Mathematics
University of California - Berkeley
plamen@math.berkeley.edu

August, 1999

Supported by NSF and DOE
INTRODUCTION

• Structured Matrices
  – Dense
  – Only depend on \( O(n) \) parameters
  – Want \( LU \), inverse and solution of \( Ax = b \) faster than \( O(n^3) \) and Matrix-vector multiply faster than \( O(n^2) \).

• Examples:
  – Cauchy \( C_{ij} = \frac{1}{x_i - y_j} \), par. \( x_1, ..., x_n, y_1, ..., y_n \).
  – Vandermonde \( V_{ij} = x_i^{j-1} \)
  – Toeplitz \( T_{ij} = t_{i-j} \)

• Inverses of these Matrices do not have same structure but still ”structured” in some sense.

• Toeplitz matrix

\[
\begin{bmatrix}
t_0 & t_1 & t_2 & \cdots & t_{n-1} \\
t_{-1} & t_0 & t_1 & \cdots & t_{n-2} \\
t_{-2} & t_{-1} & t_0 & \cdots & t_{n-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
t_{-n+1} & t_{-n+2} & t_{-n+3} & \cdots & t_0
\end{bmatrix}
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & t_0 & t_1 & \cdots & t_{n-2} \\
0 & t_{-1} & t_0 & \cdots & t_{n-3} \\
0 & \vdots & \vdots & \ddots & \vdots \\
0 & t_{-n+2} & t_{-n+3} & \cdots & t_0
\end{bmatrix}
= 
\begin{bmatrix}
t_0 & t_1 & t_2 & \cdots & t_{n-1} \\
t_{-1} & 0 & 0 & \cdots & 0 \\
t_{-2} & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
t_{-n+1} & 0 & 0 & \cdots & 0
\end{bmatrix}
\]

• We can recover \( T \) from the Right Hand Side

• \( T - Z_0 \cdot T \cdot Z_0^T = \text{Rank-2-Matrix} \), where

\[
Z_\phi = 
\begin{bmatrix}
0 & 0 & 0 & 0 & \phi \\
-1 & 0 & 0 & \cdots & 0 \\
0 & -1 & 0 & \cdots & 0 \\
0 & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 1 & 0
\end{bmatrix}
\]

• Better yet to write \( Z_1 \cdot T - T \cdot Z_1 = G \cdot B \), because \( Z_1 \) - diagonalizable.
Definition of Displacement Structure

• Given a matrix $R$ we have displacement operator

$$\nabla_{F,A}(R) = FR - RA$$

if $\nabla_{F,A}(R)$ has low rank compared to $n$ we say that $R$ has low displacement rank with respect to the operator $\nabla_{F,A}$ and we write

$$\nabla_{F,A}(R) = FR - RA = GB$$

where $G$ is $n \times \alpha$ and $B$ is $\alpha \times n$.

• For Toeplitz matrix $T$ we have $\text{rank}(\nabla_{Z_1,Z_{-1}}(R)) = 2$

• All matrices with low $\{Z_1, Z_{-1}\}$-rank are called TOEPLITZ-LIKE

• Many possible choices of $F, A$ for a particular matrix, only few useful:

  - Need to be able to recover each entry of $R$ from $F, A, G, B$ in $O(1)$ operations per entry
  - Either $F$ and $A^*$ - lower triangular, or diagonalizable (true for all basic classes)
Basic classes of Structured Matrices

Toeplitz-like: \( F = Z_1 \quad A = Z_{-1} \)

Toeplitz-plus-Hankel-like: \( F = Y_{00} \quad A = Y_{11} \)

Cauchy-like: \( F = \text{diag}(c_1, ..., c_n) \quad A = \text{diag}(d_1, ..., d_n) \)

Vandermonde-like: \( F = \text{diag}(\frac{1}{x_1}, ..., \frac{1}{x_n}) \quad A = Z_1 \)

Chebyshev-Vandermonde-like: \( F = \text{diag}(x_1, ..., x_n) \quad A = Y_{\gamma,\delta} \)

Polynomial-Vandermonde-like: \( F = \text{diag}(x_1, ..., x_n) \quad A = \text{ConfederateMatrix} \)

where

\[
Z_\phi = \begin{bmatrix}
0 & 0 & \cdots & 0 & \phi \\
1 & 0 & \cdots & \cdots & 0 \\
0 & 1 & \cdots & \cdots & \cdots \\
\vdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & 0 & 1 & 0 \\
\end{bmatrix}, \quad *Y_{\gamma,\delta} = \begin{bmatrix}
\gamma & 1 & 0 & \cdots & 0 \\
1 & 0 & 1 & \cdots & \cdots \\
0 & 1 & \cdots & \cdots & 0 \\
\vdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & 0 & 1 & \delta \\
\end{bmatrix},
\]

i.e. \( Z_\phi \) is the lower shift \( \phi \)-circulant matrix and \( Y_{\gamma,\delta} = Z_0 + Z_0^T + \gamma e_1 e_1^T + \delta e_1 e_1^T \).
Displacement Structure is Inherited During

- Inversion $FR - RA = GB$ implies $AR^{-1} - R^{-1}F = (-R^{-1}G)(BR^{-1})$
- Similarity transformations on $F, A$:
  \[
  T_{1}F_{1}T_{1}^{-1}R - RT_{2}A_{2}T_{2}^{-1} = GB
  \]
  implies
  \[
  F_{1}(T_{1}^{-1}RT_{2}) - (T_{1}^{-1}RT_{2})A_{2} = (T_{1}^{-1}G)(BT_{2})
  \]
- Schur Complementation

Let $R = \begin{bmatrix} d_{1} & u_{1} \\ l_{1} & R_{22}^{(1)} \end{bmatrix}$ satisfy

\[
\nabla_{F_{1},A_{1}}(R_{1}) = \begin{bmatrix} f_{1} & 0 \\ F_{2} \end{bmatrix} \cdot R_{1} - R_{1} \cdot \begin{bmatrix} a_{1} & * \\ 0 & A_{2} \end{bmatrix} = G_{1} \cdot B_{1}
\]

$(G_{1} \in C^{m \times \alpha}, B \in C^{\alpha \times n})$.

If $d_{1} \neq 0$ the Schur complement $R_{2} = R_{22}^{(1)} - \frac{l_{1}u_{1}}{d_{1}}$ satisfies the displacement equation

\[
F_{2} \cdot R_{2} - R_{2} \cdot A_{2} = G_{2} \cdot B_{2},
\]

where

\[
\begin{bmatrix} 0 \\ G_{2} \end{bmatrix} = G_{1} - \begin{bmatrix} 1 \\ \frac{1}{d_{1}}l_{1} \end{bmatrix} \cdot g_{1}, \quad \begin{bmatrix} 0 & B_{2} \end{bmatrix} = B_{1} - b_{1} \cdot \begin{bmatrix} 1 & \frac{1}{d_{1}}u_{1} \end{bmatrix},
\]

where $g_{1}$ and $b_{1}$ are the first row of $G_{1}$ and the first column of $B_{1}$, respectively.

- Pivoting, if $F$ is diagonal. Let $FR - RA = GB, P$-perm.

  \[
  (PFP^{T})(PR) - (PR)A = (PG)B
  \]

$PFP^{T}$ is still diagonal and pivoting translates into swapping 2 diagonal entries of $F$ and 2 rows of $G$. 
Fast $O(n^2)$ inversion

We can solve $Rx = b$ in $O(n^2)$ time if $FR - RA = GB$. The inverse satisfies

$$AR^{-1} - R^{-1}F = -(R^{-1}G)(BR^{-1})$$

- Solve $Rx = g_i$ where $g_i$ are the columns of $G$. This is $R^{-1}G$
- Solve $R^Tx = b_i^T$ where $b_i$ are the rows of $B$. The solutions are $R^{-T}B^T$. Transpose and we get $BR^{-1}$.
- Then we know the $\{A, F\}$-generators of $R^{-1}$ and can recover $R^{-1}$ in $O(n^2)$ time
- Total cost =
  - Solving $2\alpha$ linear equations with $\alpha$-small = $O(n^2)$.
  - $O(n^2)$ for recovery of $R^{-1}$ from its displacement equation.
  - Total Cost = $O(n^2)$.
Other Classes of Structured Matrices

- Toeplitz-like: \( F = Z_1 \) \( A = Z_{-1} \)
- Toeplitz-plus-Hankel-like: \( F = Y_{00} \) \( A = Y_{11} \)
- Cauchy-like: \( F = \text{diag}(c_1, \ldots, c_n) \) \( A = \text{diag}(d_1, \ldots, d_n) \)
- Vandermonde-like: \( F = \text{diag}(\frac{1}{x_1}, \ldots, \frac{1}{x_n}) \) \( A = Z_1 \)
- Chebyshev-Vandermonde-like: \( F = \text{diag}(x_1, \ldots, x_n) \) \( A = \text{ConfederateMatrix} \)

where

\[
Z_\phi = \begin{bmatrix}
0 & 0 & \cdots & 0 & \phi \\
1 & 0 & \cdots & \cdots & 0 \\
0 & 1 & \cdots & \cdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 1 & 0
\end{bmatrix}, \quad Y_{\gamma,\delta} = \begin{bmatrix}
\gamma & 1 & 0 & \cdots & 0 \\
1 & 0 & 1 & \cdots & \vdots \\
0 & 1 & \cdots & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 1 & \delta
\end{bmatrix},
\]

i.e. \( Z_\phi \) is the lower shift \( \phi \)-circulant matrix and \( Y_{\gamma,\delta} = Z_0 + Z_0^T + \gamma e_1 e_1^T + \delta e_1 e_1^T \).

- All \( F \)’s and \( A \)’s diagonal or diagonalizable (using Fast Trigonometric Transforms, diagonal scaling and products thereof.)
- We can transform \( R \) into Cauchy-like matrix
- Then we can apply Fast GEPP
- If \( ARB \) is Cauchy-like and \( ARB = PLU \) then \( R = A^{-1}PLUB^{-1} \).
- Applying \( A \) and \( B \) to the right hand side of \( Rx = b \) costs \( \leq O(n \log n) \)
- Overall cost of solving \( Rx = b \) is still \( O(n^2) \).