Structured Matrices

Structured $m \times n$ matrices include those that

- have fewer than $mn$ degrees of freedom

- the $k$ degrees of freedom result in an organization of the zero/nonzero structure or a factorization that allows the important operations to be done in fewer computations than would normally be expected with an arbitrary $m \times n$ matrix
For example, one or more of the following may be done faster

- matrix-vector or matrix-matrix multiplication
- solving systems if $m = n$
- least squares if $m \neq n$
- eigenvalues and/or eigenvectors
- singular value decomposition
- evaluating functions of the matrix
Discrete Fourier Transform

Reading:

- Computational Frameworks for the Fast Fourier Transform, C. Van Loan, SIAM
- The DFT: An Owner’s Manual for the Discrete Fourier Transform, Briggs and Henson, SIAM
- GV96 Chapter 4

The discrete Fourier transform of a vector $x \in \mathbb{C}^n$ can be defined via the application of a matrix $F \in \mathbb{C}^{n \times n}$

$$f = Fx$$
• Let $\omega^k \in \mathbb{C}, \ k = 0, \ldots, n - 1$ be the $n$th roots of unity where $\omega = e^{j2\pi/n}$ and $j = \sqrt{-1}$.

• Let an overline, e.g., $\overline{A}$, represent complex conjugation.
\[ F^H = \frac{1}{\sqrt{n}} \begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\
1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)(n-1)}
\end{pmatrix} \]
Image processing often makes use of a two-dimensional transform. The two-dimensional transform of an image represented in a matrix $A$ is

$$FAF^T$$

So $FAe_k$ is the discrete Fourier transform of the $k$-th column of the image and $e_i^T A F^T$ is the discrete Fourier transform of the $i$-th row of the image.
Useful Properties

- The matrix $F^H$ is symmetric, i.e., $(F^H)^T = F^H$.

- $F$ is created simply by replacing all $\omega^k$ with $\overline{\omega^k} = \overline{\omega}^k$, i.e., $F^H = \overline{F}$.

- $F$ is also symmetric, i.e., $F = F^T$.

- $F$ is unitary
The proof that $F$ is unitary helps introduce some of the interactions of the basic properties of the DFT.

- define $P(\gamma) = \sum_{i=0}^{n-1} \gamma^i$ and note that by definition $(\omega^k)^n - 1 = 0$.

- If $\rho$ is a root of of $P(\gamma)$ then

\[
\begin{align*}
P(\rho) &= 0 \\
\sum_{i=1}^{n-1} \rho^i &= -1 \\
\rho P(\rho) - \rho^n &= -1 \\
\rho P(\rho) &= \rho^n - 1
\end{align*}
\]

- Therefore, $\omega^k P(\omega^k) = (\omega^k)^n - 1 = 0$

- Since $|\omega^k| = 1$, it follows that

\[
P(\omega^k) = \sum_{i=0}^{n-1} (\omega^k)^i = 0.
\]
To see that $F^H F = I$ consider first the diagonal elements $\alpha_{ii} = e_i^H F^H F e_i$,:

$$\alpha_{ii} = \frac{1}{n} \left( (\omega^{i-1})^0 (\omega^{i-1})^1 \ldots (\omega^{i-1})^{n-1} \right) \left( \begin{array}{c} (\omega^{i-1})^0 \\ (\omega^{i-1})^1 \\ \vdots \\ (\omega^{i-1})^{n-1} \end{array} \right) = 1$$
Since the $I$ is Hermitian we need only consider the lower triangular elements, $\alpha_{ki} = e_k^H F^H F e_i$, with $k > i$. We have

$$\alpha_{ki} = \frac{1}{n} \left( (\omega^k - 1)^0 \quad (\omega^k - 1)^1 \quad \ldots \quad (\omega^k - 1)^{n-1} \right) \begin{pmatrix} \frac{(\omega^i - 1)^0}{(\omega^i - 1)^1} \\ \vdots \\ \frac{(\omega^i - 1)^{n-1}}{(\omega^i - 1)^{n-1}} \end{pmatrix}$$

$$= (\omega^k - 1)^0 \frac{(\omega^i - 1)^0}{(\omega^i - 1)^1} + \cdots + (\omega^k - 1)^{n-1} \frac{(\omega^i - 1)^{n-1}}{(\omega^i - 1)^{n-1}}$$

$$= (\omega^k - 1)^i \frac{(\omega^i - 1)^i}{(\omega^i - 1)^1} + \cdots + (\omega^k - 1)^{i-1} \frac{(\omega^i - 1)^{i-1}}{(\omega^i - 1)^{i-1}}$$

$$= (\omega^k - 1)^i + \cdots + (\omega^k - 1)^{i-1}$$

$$= P(\omega^k - i)$$

$$= 0$$
A final property of interest relates $F$ and $F^H$. The discrete Fourier transform matrix and its inverse are related by a simple column permutation, i.e., there exists a permutation matrix $P$ (an identity matrix with its columns or rows interchanged) such that $FP = F^H$. The permutation matrix is

$$P = \begin{pmatrix} e_1 & e_n & e_{n-1} & \cdots & e_2 \end{pmatrix}$$

To see this note that there is a simple relationship between a root of unity its conjugate, and its powers, $\omega^m = \omega^{m \mod n}$ and $\overline{\omega^k} = \omega^{n-k}$ where $0 \leq k \leq n - 1$. As a result the columns pair up as conjugate pairs, the $i$th column pairs with column $(-(i-1) \mod n) + 1$. Note also that for any permutation matrix $P^{-1} = P^H = P^T$. 
• The DFT is usually used to put a science or engineering problem into the “frequency” or “transform” domain.

• This often yields a simpler form of the problem or at least one that allows more intuitive thought (for the engineer or scientist)

• Essentially, it is a change of coordinates

• It can also be viewed as computing the coefficients of the vector $x$ relative to the Fourier basis given by the columns of $F$.

• **Most importantly there is a Fast Fourier Transform.**

• The FFT can be derived in many ways (see the text for a matrix form).

• It requires $O(n \log n)$ computations rather than $O(n^2)$. 
So for the Fourier Transform we have

- Fast matrix-vector and matrix-matrix products $Fv$ and $FA$

- Fast solutions to $Fv = b$ via $v = F^Hb$

- Fast projections to the “frequency” domain and simple truncation for approximation. (Band pass filtering)

- A fast component to more complicated signal and image processing algorithms
DEFINITION: An \( m \times n \) Toeplitz matrix has constant diagonal values, i.e., \( \tau_{ij} = \tau_{i+1,j+1} \)

For example for \( n = m = 4 \)

\[
\begin{pmatrix}
a & b & c & d \\
e & a & b & c \\
f & e & a & b \\
g & f & e & a \\
\end{pmatrix}
\]

is a nonhermitian Toeplitz matrix.
**DEFINITION:** An $m \times n$ circulant matrix is a Toeplitz matrix for which the $i + 1$-st row is a circular shift of the $i$-th row.

For example for $n = m = 4$

$$
\begin{pmatrix}
a & b & c & d \\
d & a & b & c \\
c & d & a & b \\
b & c & d & a
\end{pmatrix}
$$

is a nonhermitian circulant matrix.
The $n$ parameters that define an $n \times n$ circulant matrix allows an alternative definition of a circulant matrix.

A circulant matrix can be defined by a vector $c \in \mathbb{C}^n$ and the circulant shift matrix $Z = (e_n \ e_1 \ e_2 \ \cdots \ e_{n-1})$.

We have a circulant matrix

$$C = \begin{bmatrix}
    c^T \\
    c^T Z \\
    \vdots \\
    c^T Z^{n-1}
\end{bmatrix}.$$
The form of $C$ can be expressed in terms of polynomials as well. Let $e_k^T c = \gamma_{k-1}$. Then notice that

$$C = \sum_{k=0}^{n-1} \gamma_k Z^k$$

For our example

$$C' = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b & 0 & 0 \\ 0 & 0 & b & 0 \\ 0 & 0 & 0 & b \\ b & 0 & 0 & 0 \end{pmatrix}$$

$$+ \begin{pmatrix} 0 & 0 & c & 0 \\ 0 & 0 & 0 & c \\ c & 0 & 0 & 0 \\ 0 & c & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & d \\ d & 0 & 0 & 0 \\ 0 & d & 0 & 0 \\ 0 & 0 & d & 0 \end{pmatrix}$$
Any circulant matrix is diagonalized by the Fourier matrix $F$, i.e., for any circulant matrix $C$ there exists a diagonal matrix $\Lambda$ such that

\[
C = F^H \Lambda F
\]

\[
C F^H = F^H \Lambda
\]

In other words, the columns of $F^H$ are the eigenvectors of $C$ and the corresponding elements of $\Lambda$ are the eigenvalues.
• Circulant matrices are closed under addition and multiplication

• The eigenvalues of the sum are the sums of the eigenvalues.

• The eigenvalues of the product are the products of the eigenvalues.

• Circulant matrices commute under multiplication.
It is obvious that $C^H$ must be a circulant matrix. But, $C^T$ is diagonalized by $F^H$ and so $C^T$ is a circulant matrix

$$
C = F^H \Lambda F \\
C^T = (F^H \Lambda F)^T \\
= F^T \Lambda (F^H)^T \\
= F \Lambda F^H \\
= FPP^H \Lambda PP^H F^H \\
= (FP)(P^H \Lambda P)(P^H F^H) \\
= F^H (P^H \Lambda P) F \\
= F^H \Gamma F
$$

Since $\Lambda$ is a diagonal matrix it follows that $\Gamma$ is also a diagonal matrix whose elements are those of $\Lambda$ permuted appropriately.
The vector \( c \) and the eigenvalues of \( C \) can be related due to the fact that \( Z \) is a circulant matrix and

\[
Z = F^H \Omega F
\]

where \( \Omega = \text{diag}(1, \omega, \ldots, \omega^{n-1}) \).

It follows that

\[
\begin{align*}
C &= \sum_{k=0}^{n-1} \gamma_k Z^k \\
F C F^H &= \sum_{k=0}^{n-1} \gamma_k F Z^k F^H \\
\Lambda &= \sum_{k=0}^{n-1} \gamma_k (F Z F^H)^k \\
\Lambda &= \sum_{k=0}^{n-1} \gamma_k (\Omega)^k
\end{align*}
\]
Therefore, if we let $e_i^H \wedge e_i = \lambda_{i-1}$ we have

$$
\lambda_i = \gamma_0 + \gamma_1(\omega^i) + \gamma_2(\omega^i)^2 + \cdots + \gamma_{n-1}(\omega^i)^{n-1}
$$

which given the definition of $F^H$ can be written as

$$
v = \sqrt{n}F^H c
$$

where $v^T = (\lambda_0 \; \lambda_1 \; \cdots \; \lambda_{n-1})$. 
• matrix-vector and matrix-matrix products are faster due to FFT.

• Solving \( Cx = b \) is easy via matrix-vector products and is faster due to FFT.

• The eigenvalue problem is is easy via matrix-vector products due to the relationship between \( v \) and \( c \) and is faster due to FFT.

• The circulant matrix-vector and matrix-matrix products give fast Toeplitz matrix-vector and matrix-matrix products.
Consider the following Toeplitz matrix and an associated circulant matrix.

\[
T = \begin{pmatrix}
\alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\
\alpha_5 & \alpha_1 & \alpha_2 & \alpha_3 \\
\alpha_6 & \alpha_5 & \alpha_1 & \alpha_2 \\
\alpha_7 & \alpha_6 & \alpha_5 & \alpha_1
\end{pmatrix}
\]
\[ C' = \begin{pmatrix}
\alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_7 & \alpha_6 & \alpha_5 \\
\alpha_5 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_7 & \alpha_6 \\
\alpha_6 & \alpha_5 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_7 \\
\alpha_7 & \alpha_6 & \alpha_5 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\
\alpha_4 & \alpha_7 & \alpha_6 & \alpha_5 & \alpha_1 & \alpha_2 & \alpha_3 \\
\alpha_3 & \alpha_4 & \alpha_7 & \alpha_6 & \alpha_5 & \alpha_1 & \alpha_2 \\
\alpha_2 & \alpha_3 & \alpha_4 & \alpha_7 & \alpha_6 & \alpha_5 & \alpha_1 \\
\end{pmatrix} \]
\[ Cv = \begin{pmatrix} T & A \\ B & E \end{pmatrix} \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} Tx \\ Bx \end{pmatrix} \]

So \( Tx \) can be found by using FFT’s of approximately length \( 2n \) via circulant matrix-vector multiplication.