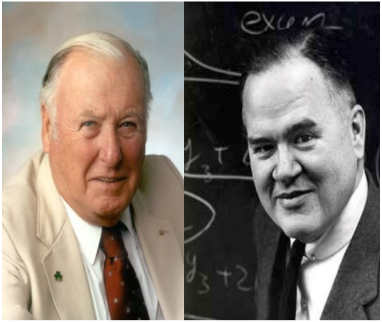
FFT History

- Cooley and Tukey FFT algorithm, 1965



James William Cooley (1926-) — John Wilder Tukey (1915-2000)

Discrete Trigonometric Least Squares

$$\min_{a_0,a_1,\cdots,a_m,b_1,\cdots,b_{m-1}} \sum_{j=0}^{2m-1}\left(y_j - \left(\frac{a_0}{2} + \sum_{k=1}^{m} a_k \cos(k\,x_j) + \sum_{k=1}^{m-1} b_k \sin(k\,x_j)\right)\right)^2.$$

Minimize least squares error,
$(x_j = \frac{j-m}{m}\pi,\ j = 0, 1, \cdots, 2m-1 :)$

$$a_k = \frac{1}{m}\sum_{j=0}^{2m-1} y_j \cos(k\,x_j), \quad k = 0, 1, \cdots, m,$$

$$b_k = \frac{1}{m}\sum_{j=0}^{2m-1} y_j \sin(k\,x_j), \quad k = 1, \cdots, m-1.$$

§8.6 Fast Fourier Transforms (I)

With $x_j = \dfrac{j-m}{m}\,\pi$, $j = 0, 1, \cdots, 2\,m-1$:

$$
\begin{aligned}
a_k &= \frac{1}{m}\sum_{j=0}^{2\,m-1} y_j \cos\left(k\,x_j\right), \quad k = 0, 1, \cdots, m, \\
b_k &= \frac{1}{m}\sum_{j=0}^{2\,m-1} y_j \sin\left(k\,x_j\right), \quad k = 1, \cdots, m-1.
\end{aligned}
$$

§8.6 Fast Fourier Transforms (I)

With $x_j = \frac{j-m}{m}\,\pi,\ j = 0, 1, \cdots, 2\,m-1$:

$$a_k = \frac{1}{m}\sum_{j=0}^{2\,m-1} y_j \cos\left(k\,x_j\right), \quad k = 0, 1, \cdots, m,$$

$$b_k = \frac{1}{m}\sum_{j=0}^{2\,m-1} y_j \sin\left(k\,x_j\right), \quad k = 1, \cdots, m-1.$$

Define $i = \sqrt{-1}$. For $k = 0, 1, \cdots, m$:

$$a_k + i\,b_k = \frac{1}{m}\sum_{j=0}^{2\,m-1} y_j \left(\cos\left(k\,x_j\right) + i \sin\left(k\,x_j\right)\right)$$

§8.6 Fast Fourier Transforms (I)

With $x_j = \frac{j-m}{m}\,\pi$, $j = 0, 1, \cdots, 2\,m-1$:

$$\begin{aligned} a_k &= \frac{1}{m}\sum_{j=0}^{2\,m-1} y_j \cos\left(k\,x_j\right), \quad k = 0, 1, \cdots, m, \\ b_k &= \frac{1}{m}\sum_{j=0}^{2\,m-1} y_j \sin\left(k\,x_j\right), \quad k = 1, \cdots, m-1. \end{aligned}$$

Define $i = \sqrt{-1}$. For $k = 0, 1, \cdots, m$:

$$\begin{aligned} a_k + i\,b_k &= \frac{1}{m}\sum_{j=0}^{2\,m-1} y_j\,\left(\cos\left(k\,x_j\right) + i\sin\left(k\,x_j\right)\right) \\ &= \frac{1}{m}\sum_{j=0}^{2\,m-1} y_j \exp\left(i\,k\,x_j\right) = \frac{1}{m}\sum_{j=0}^{2\,m-1} y_j \exp\left(i\,k\,\frac{j-m}{m}\,\pi\right) \end{aligned}$$

§8.6 Fast Fourier Transforms (I)

With $x_j = \frac{j-m}{m}\pi$, $j = 0, 1, \cdots, 2m-1$:

$$a_k = \frac{1}{m}\sum_{j=0}^{2m-1} y_j \cos(k\,x_j), \quad k = 0, 1, \cdots, m,$$

$$b_k = \frac{1}{m}\sum_{j=0}^{2m-1} y_j \sin(k\,x_j), \quad k = 1, \cdots, m-1.$$

Define $i = \sqrt{-1}$. For $k = 0, 1, \cdots, m$:

$$\begin{aligned} a_k + i\,b_k &= \frac{1}{m}\sum_{j=0}^{2m-1} y_j \left(\cos(k\,x_j) + i\sin(k\,x_j)\right) \\ &= \frac{1}{m}\sum_{j=0}^{2m-1} y_j \exp(i\,k\,x_j) = \frac{1}{m}\sum_{j=0}^{2m-1} y_j \exp\left(i\,k\,\frac{j-m}{m}\pi\right) \\ &= \frac{\exp(i\,k\,\pi)}{m}\sum_{j=0}^{2m-1} y_j \exp\left(\frac{i\,k\,j}{m}\pi\right) \stackrel{def}{=} \frac{(-1)^k}{m}c_k \end{aligned}$$

Fast Fourier Transforms (II)

$$c_k = \sum_{j=0}^{2\,m-1} y_j \, \mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \quad k = 0, 1, \cdots, 2\,m-1.$$

Fast Fourier Transforms (II)

$$c_k = \sum_{j=0}^{2m-1} y_j \, \mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \quad k = 0, 1, \cdots, 2m-1.$$

- Only $c_0, \cdots, c_m$ are needed for discrete least squares.

Fast Fourier Transforms (II)

$$c_k = \sum_{j=0}^{2\,m-1} y_j \,\mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \quad k = 0, 1, \cdots, 2\,m-1.$$

- Only $c_0, \cdots, c_m$ are needed for discrete least squares.
- $O\left(m^2\right)$ operations for computing needed coefficients.

Fast Fourier Transforms (II)

$$c_k = \sum_{j=0}^{2m-1} y_j \, \mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \quad k = 0, 1, \cdots, 2m-1.$$

- Only $c_0, \cdots, c_m$ are needed for discrete least squares.
- $O\left(m^2\right)$ operations for computing needed coefficients.
- FFT computes $c_0, \cdots, c_m, \cdots, c_{2m-1}$ in $O\left(m \, \mathbf{log} \, m\right)$ operations.

Fast Fourier Transforms (II)

$$c_k = \sum_{j=0}^{2m-1} y_j \,\mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \quad k = 0, 1, \cdots, 2\,m-1.$$

- Only $c_0, \cdots, c_m$ are needed for discrete least squares.
- $O\left(m^2\right)$ operations for computing needed coefficients.
- FFT computes $c_0, \cdots, c_m, \cdots, c_{2\,m-1}$ in $O\left(m\,\mathbf{log}\,m\right)$ operations.
- $c_{m+1}, \cdots, c_{2\,m-1}$ are computed only as part of FFT.

Fast Fourier Transforms (II)

$$c_k = \sum_{j=0}^{2\,m-1} y_j \, \mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \quad k = 0, 1, \cdots, 2\,m-1.$$

- Only $c_0, \cdots, c_m$ are needed for discrete least squares.
- $O\left(m^2\right)$ operations for computing needed coefficients.
- FFT computes $c_0, \cdots, c_m, \cdots, c_{2\,m-1}$ in $O\left(m\,\mathbf{log}\,m\right)$ operations.
- $c_{m+1}, \cdots, c_{2\,m-1}$ are computed only as part of FFT.
- FFT only works when m is power of 2 $\left(m = 2^p.\right)$

For any given vector $\mathbf{y} = (y_0, y_1, \cdots, y_{2m-1})^T$, let $\texttt{FFT}(\mathbf{y})$ be the RECURSIVE FAST FOURIER TRANSFORM

$$\begin{aligned} \mathbf{c} &\stackrel{def}{=} \texttt{FFT}(\mathbf{y}) \stackrel{def}{=} (c_0, c_1, \cdots, c_{2m-1})^T, \quad \text{where} \\ c_k &= \sum_{j=0}^{2m-1} y_j \, \mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \quad k = 0, 1, \cdots, 2m-1. \end{aligned}$$

For any given vector $\mathbf{y} = (y_0, y_1, \cdots, y_{2m-1})^T$, let $\texttt{FFT}(\mathbf{y})$ be the RECURSIVE FAST FOURIER TRANSFORM

$$\begin{aligned} \mathbf{c} &\stackrel{def}{=} \texttt{FFT}(\mathbf{y}) \stackrel{def}{=} (c_0, c_1, \cdots, c_{2m-1})^T, \quad \text{where} \\ c_k &= \sum_{j=0}^{2m-1} y_j \, \mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \quad k = 0, 1, \cdots, 2m-1. \end{aligned}$$

- Assume $m = 2M$ is even.

For any given vector $\mathbf{y} = (y_0, y_1, \cdots, y_{2m-1})^T$, let $\texttt{FFT}(\mathbf{y})$ be the RECURSIVE FAST FOURIER TRANSFORM

$$\begin{aligned} \mathbf{c} &\stackrel{def}{=} \texttt{FFT}(\mathbf{y}) \stackrel{def}{=} (c_0, c_1, \cdots, c_{2m-1})^T, \quad \text{where} \\ c_k &= \sum_{j=0}^{2m-1} y_j \, \mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \quad k = 0, 1, \cdots, 2m-1. \end{aligned}$$

- Assume $m = 2M$ is even.
- Will recursively compute $\texttt{FFT}(\mathbf{y})$ via $\texttt{FFT}(\mathbf{y}_{\mathbf{even}})$ and $\texttt{FFT}(\mathbf{y}_{\mathbf{odd}})$, where

$$\mathbf{y}_{\mathbf{even}} = (y_0, y_2, \cdots, y_{2m-2})^T, \quad \mathbf{y}_{\mathbf{odd}} = (y_1, y_3, \cdots, y_{2m-1})^T.$$

Recursive Fast Fourier Transform for $m = 2\,M$ (I)

$$\texttt{FFT}\,(\mathbf{y}) \stackrel{def}{=} (c_0, c_1, \cdots, c_{2\,m-1})^T, c_k = \sum_{j=0}^{2\,m-1} y_j \,\mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \; k = 0, 1, \cdots, 2m-1.$$

Recursive Fast Fourier Transform for $m = 2\,M$ (I)

$$\mathtt{FFT}(\mathbf{y}) \stackrel{def}{=} (c_0, c_1, \cdots, c_{2\,m-1})^T, \; c_k = \sum_{j=0}^{2\,m-1} y_j \,\mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \; k = 0, 1, \cdots, 2m{-}1.$$

$$c_k = \sum_{j=0}^{2\,M-1} y_{2\,j}\,\mathbf{exp}\left(\frac{i\,k\,(2\,j)}{2\,M}\,\pi\right) + \sum_{j=0}^{2\,M-1} y_{2\,j+1}\,\mathbf{exp}\left(\frac{i\,k\,(2\,j+1)}{2\,M}\,\pi\right)$$

Recursive Fast Fourier Transform for $m = 2\,M$ (I)

$$\texttt{FFT}\,(\mathbf{y}) \overset{def}{=} (c_0, c_1, \cdots, c_{2\,m-1})^T,\ c_k = \sum_{j=0}^{2\,m-1} y_j \,\mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right),\ k = 0, 1, \cdots, 2m\text{–}1.$$

$$\begin{aligned} c_k &= \sum_{j=0}^{2\,M-1} y_{2\,j}\,\mathbf{exp}\left(\frac{i\,k\,(2\,j)}{2\,M}\,\pi\right) + \sum_{j=0}^{2\,M-1} y_{2\,j+1}\,\mathbf{exp}\left(\frac{i\,k\,(2\,j+1)}{2\,M}\,\pi\right) \\ &= \sum_{j=0}^{2\,M-1} y_{2\,j}\,\mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right) + \mathbf{exp}\left(\frac{i\,k}{2\,M}\,\pi\right) \sum_{j=0}^{2\,M-1} y_{2\,j+1}\,\mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right) \end{aligned}$$

Recursive Fast Fourier Transform for $m = 2\,M$ (I)

$$\texttt{FFT}\,(\mathbf{y}) \stackrel{def}{=} (c_0, c_1, \cdots, c_{2\,m-1})^T, c_k = \sum_{j=0}^{2\,m-1} y_j\, \mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right),\ k = 0, 1, \cdots, 2m\text{–}1.$$

$$\begin{aligned}
c_k &= \sum_{j=0}^{2\,M-1} y_{2\,j}\, \mathbf{exp}\left(\frac{i\,k\,(2\,j)}{2\,M}\,\pi\right) + \sum_{j=0}^{2\,M-1} y_{2\,j+1}\, \mathbf{exp}\left(\frac{i\,k\,(2\,j+1)}{2\,M}\,\pi\right) \\
&= \sum_{j=0}^{2\,M-1} y_{2\,j}\, \mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right) + \mathbf{exp}\left(\frac{i\,k}{2\,M}\,\pi\right) \sum_{j=0}^{2\,M-1} y_{2\,j+1}\, \mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right) \\
c_{k+m} &= \sum_{j=0}^{2\,M-1} y_{2\,j}\, \mathbf{exp}\left(\frac{i\,(k+m)\,j}{M}\,\pi\right) \\
&\quad + \mathbf{exp}\left(\frac{i\,(k+m)}{2\,M}\,\pi\right) \sum_{j=0}^{2\,M-1} y_{2\,j+1}\, \mathbf{exp}\left(\frac{i\,(k+m)\,j}{M}\,\pi\right)
\end{aligned}$$

Recursive Fast Fourier Transform for $m = 2\,M$ (I)

$$\texttt{FFT}\,(\mathbf{y}) \stackrel{def}{=} (c_0, c_1, \cdots, c_{2\,m-1})^T, \; c_k = \sum_{j=0}^{2\,m-1} y_j \,\mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \; k = 0, 1, \cdots, 2m-1.$$

$$\begin{aligned}
c_k &= \sum_{j=0}^{2\,M-1} y_{2\,j}\,\mathbf{exp}\left(\frac{i\,k\,(2\,j)}{2\,M}\,\pi\right) + \sum_{j=0}^{2\,M-1} y_{2\,j+1}\,\mathbf{exp}\left(\frac{i\,k\,(2\,j+1)}{2\,M}\,\pi\right) \\
&= \sum_{j=0}^{2\,M-1} y_{2\,j}\,\mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right) + \mathbf{exp}\left(\frac{i\,k}{2\,M}\,\pi\right) \sum_{j=0}^{2\,M-1} y_{2\,j+1}\,\mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right) \\
c_{k+m} &= \sum_{j=0}^{2\,M-1} y_{2\,j}\,\mathbf{exp}\left(\frac{i\,(k+m)\,j}{M}\,\pi\right) \\
&\quad + \mathbf{exp}\left(\frac{i\,(k+m)}{2\,M}\,\pi\right) \sum_{j=0}^{2\,M-1} y_{2\,j+1}\,\mathbf{exp}\left(\frac{i\,(k+m)\,j}{M}\,\pi\right) \\
&= \sum_{j=0}^{2\,M-1} y_{2\,j}\,\mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right) - \mathbf{exp}\left(\frac{i\,k}{2\,M}\,\pi\right) \sum_{j=0}^{2\,M-1} y_{2\,j+1}\,\mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right)
\end{aligned}$$

Recursive Fast Fourier Transform for $m = 2\,M$ (II)

$$\texttt{FFT}\,(\mathbf{y}) \stackrel{def}{=} (c_0, c_1, \cdots, c_{2\,m-1})^T, \; c_k = \sum_{j=0}^{2\,m-1} y_j \,\mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \; k = 0, 1, \cdots, 2m{-}1.$$

Recursive Fast Fourier Transform for $m = 2\,M$ (II)

$$\texttt{FFT}\,(\mathbf{y}) \stackrel{def}{=} (c_0, c_1, \cdots, c_{2\,m-1}),^T c_k = \sum_{j=0}^{2\,m-1} y_j \,\mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \; k = 0, 1, \cdots, 2m{-}1.$$

$$c_k = \sum_{j=0}^{2\,M-1} y_{2\,j}\,\mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right) + \mathbf{exp}\left(\frac{i\,k}{2\,M}\,\pi\right) \sum_{j=0}^{2\,M-1} y_{2\,j+1}\,\mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right)$$

Recursive Fast Fourier Transform for $m = 2\,M$ (II)

$$\mathtt{FFT}\,(\mathbf{y}) \stackrel{def}{=} (c_0, c_1, \cdots, c_{2\,m-1})^T, \; c_k = \sum_{j=0}^{2\,m-1} y_j \,\mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \; k = 0, 1, \cdots, 2m{-}1.$$

$$\begin{aligned} c_k &= \sum_{j=0}^{2\,M-1} y_{2\,j}\,\mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right) + \mathbf{exp}\left(\frac{i\,k}{2\,M}\,\pi\right) \sum_{j=0}^{2\,M-1} y_{2\,j+1}\,\mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right) \\ c_{k+m} &= \sum_{j=0}^{2\,M-1} y_{2\,j}\,\mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right) - \mathbf{exp}\left(\frac{i\,k}{2\,M}\,\pi\right) \sum_{j=0}^{2\,M-1} y_{2\,j+1}\,\mathbf{exp}\left(\frac{i\,k\,j}{M}\,\pi\right) \end{aligned}$$

- First sums only involve `FFT` of $\mathbf{y}_{\mathbf{even}} = (y_0, y_2, \cdots, y_{2\,m-2})^T$,
- Second sums only involve `FFT` of $\mathbf{y}_{\mathbf{odd}} = (y_1, y_3, \cdots, y_{2\,m-1})^T$.

Recursive Fast Fourier Transform for $m = 2\,M$ (II)

$$\mathtt{FFT}\,(\mathbf{y}) \stackrel{def}{=} (c_0, c_1, \cdots, c_{2\,m-1})^T, c_k = \sum_{j=0}^{2\,m-1} y_j \,\mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \; k = 0, 1, \cdots, 2m{-}1.$$

Recursive Fast Fourier Transform for $m = 2\,M$ (II)

$$\mathtt{FFT}\,(\mathbf{y}) \stackrel{def}{=} (c_0, c_1, \cdots, c_{2\,m-1})^T, \; c_k = \sum_{j=0}^{2\,m-1} y_j \,\mathbf{exp}\left(\frac{i\,k\,j}{m}\,\pi\right), \; k = 0, 1, \cdots, 2m\text{–}1.$$

$$\mathtt{FFT}\,(\mathbf{y}) = \begin{pmatrix} c_0 \\ \vdots \\ c_{2\,M-1} \\ c_{0+m} \\ \vdots \\ c_{2\,M-1+m} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} c_0 \\ \vdots \\ c_{2\,M-1} \end{pmatrix} \\ \begin{pmatrix} c_{0+m} \\ \vdots \\ c_{2\,M-1+m} \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} \mathtt{FFT}\,(\mathbf{y}_{\mathbf{even}}) + \Delta .* \mathtt{FFT}\,(\mathbf{y}_{\mathbf{odd}}) \\ \mathtt{FFT}\,(\mathbf{y}_{\mathbf{even}}) - \Delta .* \mathtt{FFT}\,(\mathbf{y}_{\mathbf{odd}}) \end{pmatrix},$$

$$\text{where} \quad \Delta = \left(1, \mathbf{exp}\left(\frac{i}{2\,M}\,\pi\right), \cdots, \mathbf{exp}\left(\frac{i\,(2\,M-1)}{2\,M}\,\pi\right)\right)^T.$$

Fast Fourier Transforms (FFT)

Algorithm 1 `FFT`($\mathbf{y}$)

Input: vector $\mathbf{y} = (y_0, y_1, \cdots, y_{2m-1})^T$ with $m = 2^p$.
Output: Fourier Transform $\mathbf{c} = \mathtt{FFT}(\mathbf{y})$.
Algorithm:
Set $\mathbf{y_{even}} = (y_0, y_2, \cdots, y_{2m-2})^T$, $\mathbf{y_{odd}} = (y_1, y_3, \cdots, y_{2m-1})^T$.
Recursively compute $\mathbf{c_{even}} = \mathtt{FFT}(\mathbf{y_{even}})$, $\mathbf{c_{odd}} = \mathtt{FFT}(\mathbf{y_{odd}})$.
Set $M = m/2$ and compute

$$\Delta = \left(1, \mathbf{exp}\left(\frac{i}{2M}\pi\right), \cdots, \mathbf{exp}\left(\frac{i(2M-1)}{2M}\pi\right)\right)^T.$$

Compute

$$\mathtt{FFT}(\mathbf{y}) = \begin{pmatrix} \mathbf{c_{even}} + \Delta .* \mathbf{c_{odd}} \\ \mathbf{c_{even}} - \Delta .* \mathbf{c_{odd}} \end{pmatrix}.$$

Cost for Recursive Fast Fourier Transform

$$\mathtt{FFT}(\mathbf{y}) = \begin{pmatrix} \mathtt{FFT}(\mathbf{y}_{\mathbf{even}}) + \Delta. * \mathtt{FFT}(\mathbf{y}_{\mathbf{odd}}) \\ \mathtt{FFT}(\mathbf{y}_{\mathbf{even}}) - \Delta. * \mathtt{FFT}(\mathbf{y}_{\mathbf{odd}}) \end{pmatrix},$$

$$\text{where} \quad \Delta = \left(1, \mathbf{exp}\left(\frac{i}{2\,M}\,\pi\right), \cdots, \mathbf{exp}\left(\frac{i\,(2\,M-1)}{2\,M}\,\pi\right)\right)^T.$$

Let T_m be the computational cost for $\mathtt{FFT}(\mathbf{y})$:

$$T_m = 2\,T_{m/2} + K\,m, \quad \text{where}$$

- Cost of $2\,T_{m/2}$ for computing $\mathtt{FFT}(\mathbf{y}_{\mathbf{even}})$ and $\mathtt{FFT}(\mathbf{y}_{\mathbf{odd}})$.
- $O(m)$ cost for computing Δ.
- $O(m)$ cost for computing $\mathtt{FFT}(\mathbf{y})$ with $\mathtt{FFT}(\mathbf{y}_{\mathbf{even}})$, $\mathtt{FFT}(\mathbf{y}_{\mathbf{odd}})$, and Δ.

```
>> n = 1024*4; y = randn(n,1); I  = eye(n);
>> F = fft(I);
>> tic, z = F*y; toc
Elapsed time is 0.502651 seconds.
>> tic, w = fft(y); toc
Elapsed time is 0.00698113 seconds.
>> norm(w-z)/norm(w)
ans =    2.0974e-15
```

§9.1 Eigenvalues: definition and computation

- Given $n \times n$ matrix A,
- Characteristic polynomial $p(\lambda) \stackrel{def}{=} \mathbf{det}(A - \lambda I)$.
- Eigenvalues: roots of $p(\lambda) = 0$

§9.1 Eigenvalues: definition and computation

- Given $n \times n$ matrix A,
- Characteristic polynomial $p(\lambda) \stackrel{def}{=} \mathbf{det}(A - \lambda I)$.
- Eigenvalues: roots of $p(\lambda) = 0$

Geršgorin Circle Thm: Let A be an $n \times n$ matrix, and let $\mathcal{R}_i$ denote the circle in the COMPLEX PLANE with center a_{ii} and radius $\displaystyle\sum_{j=1, j \neq i}^{n} |a_{i,j}|$:

$$\mathcal{R}_i \stackrel{def}{=} \left\{ z \in \mathcal{C} \;\middle|\; |z - a_{ii}| \leq \sum_{j=1, j \neq i}^{n} |a_{i,j}| \right\}, \text{ where } \mathcal{C} = \text{complex plane.}$$

§9.1 Eigenvalues: definition and computation

- Given $n \times n$ matrix A,
- Characteristic polynomial $p(\lambda) \stackrel{def}{=} \mathbf{det}(A - \lambda I)$.
- Eigenvalues: roots of $p(\lambda) = 0$

Geršgorin Circle Thm: Let A be an $n \times n$ matrix, and let $\mathcal{R}_i$ denote the circle in the COMPLEX PLANE with center a_{ii} and radius $\sum_{j=1, j\neq i}^{n} |a_{i,j}|$:

$$\mathcal{R}_i \stackrel{def}{=} \left\{ z \in \mathcal{C} \,\middle|\, |z - a_{ii}| \leq \sum_{j=1, j\neq i}^{n} |a_{i,j}| \right\}, \text{ where } \mathcal{C} = \text{complex plane.}$$

- Eigenvalues of A are contained within $\mathcal{R} = \cup_{i=1}^{n} \mathcal{R}_i$.
- The union of any k of the circles that do not intersect with the remaining $n - k$ contains precisely k (counting multiplicities) of the eigenvalues.

Geršgorin Circles

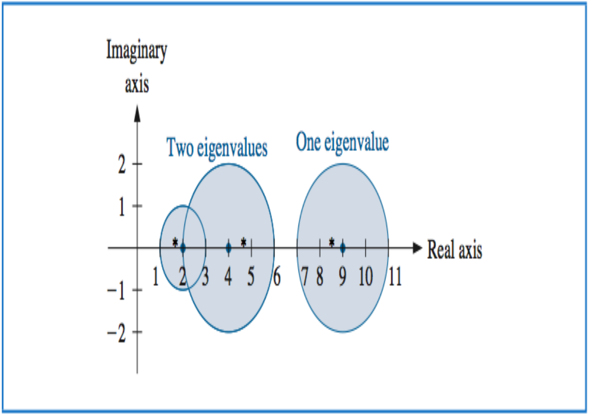
- $A = \begin{pmatrix} 4 & 1 & 1 \\ 0 & 2 & 1 \\ -2 & 0 & 9 \end{pmatrix}$, with eigenvalues

$$\lambda_1 = 1.88, \quad \lambda_2 = 4.63, \quad \lambda_3 = 8.48.$$

Geršgorin Circles

- $A = \begin{pmatrix} 4 & 1 & 1 \\ 0 & 2 & 1 \\ -2 & 0 & 9 \end{pmatrix}$, with eigenvalues

$$\lambda_1 = 1.88, \quad \lambda_2 = 4.63, \quad \lambda_3 = 8.48.$$



Proof of Geršgorin Circle Thm

- Let λ be an eigenvalue, with $\mathbf{x}$ being a (non-zero) eigenvector: $A\,\mathbf{x} = \lambda\,\mathbf{x}$.
- Write $\mathbf{x} = (x_1, \cdots, x_n)^T$ and let $|x_k| = \|\mathbf{x}\|_\infty$. The k^{th} scalar equation in $A\,\mathbf{x} = \lambda\,\mathbf{x}$ is:

$$\sum_{j=1}^{n} a_{k,j}\,x_j = \lambda\,x_k,$$

Proof of Geršgorin Circle Thm

- Let λ be an eigenvalue, with $\mathbf{x}$ being a (non-zero) eigenvector: $A\,\mathbf{x} = \lambda\,\mathbf{x}$.
- Write $\mathbf{x} = (x_1, \cdots, x_n)^T$ and let $|x_k| = \|\mathbf{x}\|_\infty$. The k^{th} scalar equation in $A\,\mathbf{x} = \lambda\,\mathbf{x}$ is:

$$\sum_{j=1}^{n} a_{k,j}\, x_j = \lambda\, x_k, \quad \text{or} \quad \sum_{j=1, j\neq k}^{n} a_{k,j}\, x_j = (\lambda - a_{k,k})\, x_k.$$

-

$$|\lambda - a_{k,k}| = \left| \frac{1}{x_k} \sum_{j=1, j\neq k}^{n} a_{k,j}\, x_j \right| \leq \sum_{j=1, j\neq k}^{n} |a_{k,j}| \frac{|x_j|}{|x_k|} \leq \sum_{j=1, j\neq k}^{n} |a_{k,j}|.$$

Proof of Geršgorin Circle Thm

- Let λ be an eigenvalue, with $\mathbf{x}$ being a (non-zero) eigenvector: $A\,\mathbf{x} = \lambda\,\mathbf{x}$.
- Write $\mathbf{x} = (x_1, \cdots, x_n)^T$ and let $|x_k| = \|\mathbf{x}\|_\infty$. The k^{th} scalar equation in $A\,\mathbf{x} = \lambda\,\mathbf{x}$ is:

$$\sum_{j=1}^{n} a_{k,j}\,x_j = \lambda\,x_k, \quad \text{or} \quad \sum_{j=1, j\neq k}^{n} a_{k,j}\,x_j = (\lambda - a_{k,k})\;x_k.$$

-

$$|\lambda - a_{k,k}| = \left| \frac{1}{x_k} \sum_{j=1, j\neq k}^{n} a_{k,j}\,x_j \right| \leq \sum_{j=1, j\neq k}^{n} |a_{k,j}| \frac{|x_j|}{|x_k|} \leq \sum_{j=1, j\neq k}^{n} |a_{k,j}|.$$

Def: Let $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}\right\}$ be a set of vectors. The set is LINEARLY INDEPENDENT if whenever

$$\alpha_1 \, \mathbf{v}^{(1)} + \alpha_2 \, \mathbf{v}^{(2)} + \cdots + \alpha_k \, \mathbf{v}^{(k)} = \mathbf{0},$$

then $\alpha_i = 0$, for each $i = 1, 2, \cdots, k$. Otherwise the set of vectors is LINEARLY DEPENDENT.

Def: Let $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}\right\}$ be a set of vectors. The set is LINEARLY INDEPENDENT if whenever

$$\alpha_1 \mathbf{v}^{(1)} + \alpha_2 \mathbf{v}^{(2)} + \cdots + \alpha_k \mathbf{v}^{(k)} = \mathbf{0},$$

then $\alpha_i = 0$, for each $i = 1, 2, \cdots, k$. Otherwise the set of vectors is LINEARLY DEPENDENT.

Thm: Let $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right\}$ be a set of linearly independent vectors in $\mathbb{R}^n$. Then for any vector $\mathbf{x} \in \mathbb{R}^n$, there is a unique collection of constants $\beta_1, \beta_2, \cdots, \beta_k$ so that

$$\mathbf{x} = \beta_1 \mathbf{v}^{(1)} + \beta_2 \mathbf{v}^{(2)} + \cdots + \beta_k \mathbf{v}^{(k)}.$$

Def: Let $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}\right\}$ be a set of vectors. The set is LINEARLY INDEPENDENT if whenever

$$\alpha_1 \mathbf{v}^{(1)} + \alpha_2 \mathbf{v}^{(2)} + \cdots + \alpha_k \mathbf{v}^{(k)} = \mathbf{0},$$

then $\alpha_i = 0$, for each $i = 1, 2, \cdots, k$. Otherwise the set of vectors is LINEARLY DEPENDENT.

Thm: Let $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right\}$ be a set of linearly independent vectors in $\mathbb{R}^n$. Then for any vector $\mathbf{x} \in \mathbb{R}^n$, there is a unique collection of constants $\beta_1, \beta_2, \cdots, \beta_k$ so that

$$\mathbf{x} = \beta_1 \mathbf{v}^{(1)} + \beta_2 \mathbf{v}^{(2)} + \cdots + \beta_k \mathbf{v}^{(k)}. \quad \text{(Set is called a BASIS.)}$$

Def: Let $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}\right\}$ be a set of vectors. The set is LINEARLY INDEPENDENT if whenever

$$\alpha_1 \mathbf{v}^{(1)} + \alpha_2 \mathbf{v}^{(2)} + \cdots + \alpha_k \mathbf{v}^{(k)} = \mathbf{0},$$

then $\alpha_i = 0$, for each $i = 1, 2, \cdots, k$. Otherwise the set of vectors is LINEARLY DEPENDENT.

Thm: Let $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right\}$ be a set of linearly independent vectors in $\mathbb{R}^n$. Then for any vector $\mathbf{x} \in \mathbb{R}^n$, there is a unique collection of constants $\beta_1, \beta_2, \cdots, \beta_k$ so that

$$\mathbf{x} = \beta_1 \mathbf{v}^{(1)} + \beta_2 \mathbf{v}^{(2)} + \cdots + \beta_k \mathbf{v}^{(k)}. \quad \text{(Set is called a BASIS.)}$$

Proof:

$$\mathbf{x} = \left(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right) \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix},$$

Def: Let $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}\right\}$ be a set of vectors. The set is LINEARLY INDEPENDENT if whenever

$$\alpha_1 \mathbf{v}^{(1)} + \alpha_2 \mathbf{v}^{(2)} + \cdots + \alpha_k \mathbf{v}^{(k)} = \mathbf{0},$$

then $\alpha_i = 0$, for each $i = 1, 2, \cdots, k$. Otherwise the set of vectors is LINEARLY DEPENDENT.

Thm: Let $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right\}$ be a set of linearly independent vectors in $\mathbb{R}^n$. Then for any vector $\mathbf{x} \in \mathbb{R}^n$, there is a unique collection of constants $\beta_1, \beta_2, \cdots, \beta_k$ so that

$$\mathbf{x} = \beta_1 \mathbf{v}^{(1)} + \beta_2 \mathbf{v}^{(2)} + \cdots + \beta_k \mathbf{v}^{(k)}. \quad \text{(Set is called a BASIS.)}$$

Proof:

$$\mathbf{x} = \left(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right) \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}, \quad \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \left(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right)^{-1} \mathbf{x}.$$

Ex: $\mathbf{v}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v}^{(2)} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$, $\mathbf{v}^{(3)} = \begin{pmatrix} 0 \\ 4 \\ 2 \end{pmatrix}$

Claim: $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}\right\}$ are LINEARLY INDEPENDENT:

$$\alpha_1\,\mathbf{v}^{(1)} + \alpha_2\,\mathbf{v}^{(2)} + \alpha_3\,\mathbf{v}^{(3)} = \mathbf{0},$$

Ex: $\mathbf{v}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{v}^{(2)} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{v}^{(3)} = \begin{pmatrix} 0 \\ 4 \\ 2 \end{pmatrix}$

Claim: $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}\right\}$ are LINEARLY INDEPENDENT:

$$\alpha_1 \mathbf{v}^{(1)} + \alpha_2 \mathbf{v}^{(2)} + \alpha_3 \mathbf{v}^{(3)} = \mathbf{0}, \quad \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 4 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \mathbf{0},$$

Ex: $\mathbf{v}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v}^{(2)} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$, $\mathbf{v}^{(3)} = \begin{pmatrix} 0 \\ 4 \\ 2 \end{pmatrix}$

Claim: $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}\right\}$ are LINEARLY INDEPENDENT:

$$\alpha_1 \mathbf{v}^{(1)} + \alpha_2 \mathbf{v}^{(2)} + \alpha_3 \mathbf{v}^{(3)} = \mathbf{0}, \quad \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 4 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \mathbf{0}, \quad \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \mathbf{0}$$

- Let $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ and let $\mathbf{x} = \beta_1 \mathbf{v}^{(1)} + \beta_2 \mathbf{v}^{(2)} + \beta_3 \mathbf{v}^{(3)}$:

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 4 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

Ex: $\mathbf{v}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{v}^{(2)} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{v}^{(3)} = \begin{pmatrix} 0 \\ 4 \\ 2 \end{pmatrix}$

Claim: $\left\{ \mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)} \right\}$ are LINEARLY INDEPENDENT:

$$\alpha_1 \mathbf{v}^{(1)} + \alpha_2 \mathbf{v}^{(2)} + \alpha_3 \mathbf{v}^{(3)} = \mathbf{0}, \ \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 4 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \mathbf{0}, \ \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \mathbf{0}$$

- Let $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ and let $\mathbf{x} = \beta_1 \mathbf{v}^{(1)} + \beta_2 \mathbf{v}^{(2)} + \beta_3 \mathbf{v}^{(3)}$:

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 4 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 4 \\ 0 & 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 + 2\,x_3 \\ 2\,x_3 - x_2 \\ \frac{1}{2}\,(x_2 - x_3) \end{pmatrix}.$$

Thm: If $A \in \mathbb{R}^{n\times n}$ is a matrix and $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct eigenvalues of A with associated eigenvectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)}$, then $\left\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)}\right\}$ is a linearly independent set.

Thm: If $A \in \mathbb{R}^{n\times n}$ is a matrix and $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct eigenvalues of A with associated eigenvectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)}$, then $\left\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)}\right\}$ is a linearly independent set.

Proof: Induction on k. Assume Thm true for $k-1$. Let

$$\alpha_1 \mathbf{x}^{(1)} + \alpha_2 \mathbf{x}^{(2)} + \cdots + \alpha_k \mathbf{x}^{(k)} = \mathbf{0}.$$

Thm: If $A \in \mathbb{R}^{n\times n}$ is a matrix and $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct eigenvalues of A with associated eigenvectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)}$, then $\left\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)}\right\}$ is a linearly independent set.

Proof: Induction on k. Assume Thm true for $k-1$. Let

$$\alpha_1\,\mathbf{x}^{(1)} + \alpha_2\,\mathbf{x}^{(2)} + \cdots + \alpha_k\,\mathbf{x}^{(k)} = \mathbf{0}.$$

Then $\alpha_1\,\lambda_k\,\mathbf{x}^{(1)} + \alpha_2\,\lambda_k\,\mathbf{x}^{(2)} + \cdots + \alpha_k\,\lambda_k\,\mathbf{x}^{(k)} = \mathbf{0}$ and

$$A\left(\alpha_1\,\mathbf{x}^{(1)} + \alpha_2\,\mathbf{x}^{(2)} + \cdots + \alpha_k\,\mathbf{x}^{(k)}\right)$$
$$= \alpha_1\,\lambda_1\,\mathbf{x}^{(1)} + \alpha_2\,\lambda_2\,\mathbf{x}^{(2)} + \cdots + \alpha_k\,\lambda_k\,\mathbf{x}^{(k)} = \mathbf{0}.$$

Thm: If $A \in \mathbb{R}^{n \times n}$ is a matrix and $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct eigenvalues of A with associated eigenvectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)}$, then $\left\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)}\right\}$ is a linearly independent set.

Proof: Induction on k. Assume Thm true for $k-1$. Let

$$\alpha_1 \mathbf{x}^{(1)} + \alpha_2 \mathbf{x}^{(2)} + \cdots + \alpha_k \mathbf{x}^{(k)} = \mathbf{0}.$$

Then $\quad \alpha_1 \lambda_k \mathbf{x}^{(1)} + \alpha_2 \lambda_k \mathbf{x}^{(2)} + \cdots + \alpha_k \lambda_k \mathbf{x}^{(k)} = \mathbf{0} \quad$ and

$$A\left(\alpha_1 \mathbf{x}^{(1)} + \alpha_2 \mathbf{x}^{(2)} + \cdots + \alpha_k \mathbf{x}^{(k)}\right)$$
$$= \alpha_1 \lambda_1 \mathbf{x}^{(1)} + \alpha_2 \lambda_2 \mathbf{x}^{(2)} + \cdots + \alpha_k \lambda_k \mathbf{x}^{(k)} = \mathbf{0}.$$

Taking difference,

$$\alpha_1 \left(\lambda_1 - \lambda_k\right) \mathbf{x}^{(1)} + \alpha_2 \left(\lambda_2 - \lambda_k\right) \mathbf{x}^{(2)} + \cdots + \alpha_{k-1} \left(\lambda_{k-1} - \lambda_k\right) \mathbf{x}^{(k-1)} = \mathbf{0}.$$

By induction, $\alpha_1 = \alpha_2 = \cdots = \alpha_{k-1} = 0$, and so $\alpha_k = 0$. $\square$

- Let $A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 2 \\ 1 & -1 & 4 \end{pmatrix}$. Eigenvalues $2, 2, 3$, eigenvectors

$$\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}.$$

- Let $A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 2 \\ 1 & -1 & 4 \end{pmatrix}$. Eigenvalues $2, 2, 3$, eigenvectors

$$\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}.$$

- Let $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. Eigenvalues $2, 2, 3$, eigenvectors

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Orthogonal Vectors

Def: A set of vectors $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}\right\}$ is called orthogonal if $\left(\mathbf{v}^{(i)}\right)^T \left(\mathbf{v}^{(j)}\right) = 0$ for all $i \neq j$.

Orthogonal Vectors

Def: A set of vectors $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}\right\}$ is called orthogonal if $\left(\mathbf{v}^{(i)}\right)^T \left(\mathbf{v}^{(j)}\right) = 0$ for all $i \neq j$. If, in addition, $\left(\mathbf{v}^{(i)}\right)^T \left(\mathbf{v}^{(i)}\right) = 1$ for all $i = 1, 2, \cdots, n$, then the set is called ORTHONORMAL.

Orthogonal Vectors

Def: A set of vectors $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}\right\}$ is called orthogonal if $\left(\mathbf{v}^{(i)}\right)^T \left(\mathbf{v}^{(j)}\right) = 0$ for all $i \neq j$. If, in addition, $\left(\mathbf{v}^{(i)}\right)^T \left(\mathbf{v}^{(i)}\right) = 1$ for all $i = 1, 2, \cdots, n$, then the set is called ORTHONORMAL.

Ex: Orthogonal set of vectors

$$\mathbf{v}^{(1)} = \begin{pmatrix} 0 \\ 4 \\ 2 \end{pmatrix}, \quad \mathbf{v}^{(2)} = \begin{pmatrix} -5 \\ -1 \\ 2 \end{pmatrix}, \quad \mathbf{v}^{(3)} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

Orthogonal Vectors

Def: A set of vectors $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}\right\}$ is called orthogonal if $\left(\mathbf{v}^{(i)}\right)^T \left(\mathbf{v}^{(j)}\right) = 0$ for all $i \neq j$. If, in addition, $\left(\mathbf{v}^{(i)}\right)^T \left(\mathbf{v}^{(i)}\right) = 1$ for all $i = 1, 2, \cdots, n$, then the set is called ORTHONORMAL.

Ex: Orthogonal set of vectors

$$\mathbf{v}^{(1)} = \begin{pmatrix} 0 \\ 4 \\ 2 \end{pmatrix}, \quad \mathbf{v}^{(2)} = \begin{pmatrix} -5 \\ -1 \\ 2 \end{pmatrix}, \quad \mathbf{v}^{(3)} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

Ex: Orthonormal set of vectors

$$\mathbf{u}^{(1)} = \frac{\mathbf{v}^{(1)}}{\left\|\mathbf{v}^{(1)}\right\|_2} = \frac{1}{2\sqrt{5}} \begin{pmatrix} 0 \\ 4 \\ 2 \end{pmatrix}, \quad \mathbf{u}^{(2)} = \frac{\mathbf{v}^{(2)}}{\left\|\mathbf{v}^{(2)}\right\|_2} = \frac{1}{\sqrt{30}} \begin{pmatrix} -5 \\ -1 \\ 2 \end{pmatrix}.$$

Thm: An orthogonal set of nonzero vectors is linearly independent.

Thm: An orthogonal set of nonzero vectors is linearly independent.

Proof: Let $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}$ be nonzero orthogonal vectors and let

$$\alpha_1 \, \mathbf{v}^{(1)} + \alpha_2 \, \mathbf{v}^{(2)} + \cdots + \alpha_k \, \mathbf{v}^{(k)} = \mathbf{0}.$$

Thm: An orthogonal set of nonzero vectors is linearly independent.

Proof: Let $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}$ be nonzero orthogonal vectors and let

$$\alpha_1 \mathbf{v}^{(1)} + \alpha_2 \mathbf{v}^{(2)} + \cdots + \alpha_k \mathbf{v}^{(k)} = \mathbf{0}.$$

Perform dot product with $\mathbf{v}^{(j)}$ on both sides,

$$\alpha_j \left(\mathbf{v}^{(j)}\right)^T \left(\mathbf{v}^{(j)}\right) = 0.$$

Therefore $\alpha_j = 0$. $\square$

Gram-Schmidt process

Thm: Let $\left\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)}\right\}$ be a set of k linearly independent vectors in $\mathbb{R}^n$. Then $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}\right\}$ defined by

$$
\begin{aligned}
\mathbf{v}^{(1)} &= \mathbf{x}^{(1)}, \\
\mathbf{v}^{(2)} &= \mathbf{x}^{(2)} - \left(\frac{\left(\mathbf{v}^{(1)}\right)^T \left(\mathbf{x}^{(2)}\right)}{\left(\mathbf{v}^{(1)}\right)^T \left(\mathbf{v}^{(1)}\right)} \right) \mathbf{v}^{(1)}, \\
\mathbf{v}^{(j)} &= \mathbf{x}^{(j)} - \sum_{i=1}^{j-1} \left(\frac{\left(\mathbf{v}^{(i)}\right)^T \left(\mathbf{x}^{(j)}\right)}{\left(\mathbf{v}^{(i)}\right)^T \left(\mathbf{v}^{(i)}\right)} \right) \mathbf{v}^{(i)}, \quad j = 3, \cdots, k,
\end{aligned}
$$

is set of k orthogonal vectors in $\mathbb{R}^n$.

Gram-Schmidt process

Thm: Let $\left\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)}\right\}$ be a set of k linearly independent vectors in $\mathbb{R}^n$. Then $\left\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(k)}\right\}$ defined by

$$
\begin{aligned}
\mathbf{v}^{(1)} &= \mathbf{x}^{(1)}, \\
\mathbf{v}^{(2)} &= \mathbf{x}^{(2)} - \left(\frac{\left(\mathbf{v}^{(1)}\right)^T \left(\mathbf{x}^{(2)}\right)}{\left(\mathbf{v}^{(1)}\right)^T \left(\mathbf{v}^{(1)}\right)} \right) \mathbf{v}^{(1)}, \\
\mathbf{v}^{(j)} &= \mathbf{x}^{(j)} - \sum_{i=1}^{j-1} \left(\frac{\left(\mathbf{v}^{(i)}\right)^T \left(\mathbf{x}^{(j)}\right)}{\left(\mathbf{v}^{(i)}\right)^T \left(\mathbf{v}^{(i)}\right)} \right) \mathbf{v}^{(i)}, \quad j = 3, \cdots, k,
\end{aligned}
$$

is set of k orthogonal vectors in $\mathbb{R}^n$.

Not much practiced in numerical computations.