TUTORIAL 4

October 10, 2024

1 Convolution (continued)

In last tutorial, we have

Proposition 1.1. Let $\mathrm{DFT}(f) = (\hat{f}_0, ..., \hat{f}_{N-1})$, we have

$$\mathrm{DFT}(f * g) = (N \cdot \hat{f}_0 \cdot \hat{g}_0, ..., N \cdot \hat{f}_{N-1} \cdot \hat{g}_{N-1})$$

Besides, we have a similar equation for a convolution in the frequency domain.

Proposition 1.2. Suppose $F = \mathrm{DFT}(f), G = \mathrm{DFT}(g)$, then

$$(F * G)(k) = \widehat{f \cdot g}(k)$$

2 Derivatives

Suppose we have nodes x_j (for $j = 0, ..., N$) in an interval $[0, 2\pi]$, and data $\{f_j\}$ from a real function $f(x)$ and wish to construct a 'global' approximation to the derivative f(x) in the interval (or at least at the nodes). Suppose f is 2π-periodic, from before, we know $p_n(x) = \sum_{k=0}^{N-1} \hat{f}(k)e^{ikx}$ is a good interpolant at the nodes $\{x_j\}$. Observe that

$$f(x) \approx p_n(x) = \sum_{k=0}^{N-1} \hat{f}(k)e^{ikx} \Rightarrow f'(x) \approx p_n'(x) = \sum_{k=0}^{N-1} ik\hat{f}(k)e^{ikx}$$

And this process can be summarized in a matrix form.

$$\left(f'(x_0), ..., f'(x_{N-1})\right)^T \approx \tilde{D}\left(f(x_0), ..., f(x_{N-1})\right)^T$$

where

$$\tilde{D} = \mathcal{F}^{-1}D\mathcal{F}, D = diag(0, i, 2i, ..., (N-1)i)$$

Note that this interpolated derivative doesn't necessarily equal the true value of $f'(x)$ and there is no theoretical guarantee about the error of $|p_n'(x_j) - f'(x_j)|$. However, in reality, we have to solve an ODE $\mathcal{L}u = f(x)$ with only $\{f(x_j)\}$ available. In such a case, all we can do is to give estimations of $\{u(x_j)\}$. Instead of using p_n and its derivatives to approximate the differential operator $\mathcal{L}$, we usually try the approach of finite difference scheme. It's also a linear interpolation method, which is using linear combinations of $u(x_j)$ to estimate $\mathcal{L}u(x_j)$. Its benefit is that the approximation error can be of some integer order r by careful design via Taylor expansion.

For example, assume $x_{k+1} = x_k + h$, when $\mathcal{L} = \frac{\partial}{\partial x}$,

$$\begin{aligned}
u'(x_j) &= \frac{u(x_{j+1}) - u(x_{j-1})}{2h} + O(h^2) \\
u'(x_j) &= \frac{3u(x_j) - 4u(x_{j-1}) + u(x_{j-2})}{2h} + O(h^2)) \\
u'(x_j) &= \frac{-u(x_{j+2}) + 8u(x_{j+1}) - 8u(x_{j-1}) + u(x_{j-2})}{12h} + O(h^4)
\end{aligned}$$

It can also apply to higher derivatives, say, $\mathcal{L} = \frac{\partial^2}{\partial x^2}$

$$\begin{aligned}
u''(x_j) &= \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{h^2} + O(h^2) \\
u''(x_j) &= \frac{2u(x_j) - 5u(x_{j-1}) + 4u(x_{j-2}) - u(x_{j-3})}{h^3} + O(h^2)) \\
u''(x_j) &= \frac{-u(x_{j+2}) + 16u(x_{j+1}) - 30u(x_j) + 16u(x_{j-1}) - u(x_{j-2})}{12h^2} + O(h^4)
\end{aligned}$$

Question 2.1. prove the above six equations.

Using a certain finite difference scheme,

$$\mathcal{L}u = f \Rightarrow D\vec{u} = \vec{f}$$

where $\vec{u} = (u(x_0), u(x_1), ..., u(x_{N-1}))^T$ and $\vec{f} = (f(x_0), f(x_1), ..., f(x_{N-1}))^T$. When x_j is near the endpoint of an interval, for example, $j = N-1$, estimating $\mathcal{L}u(x_{N-1})$ needs $x_N, x_{N+1}, ...$ which is not included in the observations $\{f_{x_j}\}$. Usually, f is assumes to be periodic, thus $x_k = x_{k+N}$. As a result, the interpolation matrix D is circulant.

Definition 2.1. (Circulant Matrix) A circulant matrix is a square matrix in which all rows are composed of the same elements and each row is rotated one element to the right relative to the preceding row. A $n \times n$ matrix C is of the form

$$\begin{bmatrix} c_0 & c_{n-1} & \cdots & c_1 \\ c_1 & c_0 & \cdots & c_2 \\ \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & \cdots & c_0 \end{bmatrix}$$

But the computation cost of directly solving this linear system is $O(N^3)$. Therefore we turn to the discrete Fourier transform for efficiency.

3 DFT and differential equation

Let $\omega = e^{+i\frac{2\pi}{N}}$ and denote the discrete Fourier transform matrix by $A(\omega)$,

$$A(\omega) = \begin{bmatrix} \omega^{0*0} & \omega^{0*1} & \cdots & \omega^{0*(N-1)} \\ \omega^{1*0} & \omega^{1*1} & \cdots & \omega^{1*(N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \omega^{(N-1)*0} & \omega^{(N-1)*1} & \cdots & \omega^{(N-1)*(N-1)} \end{bmatrix}$$

we have $\vec{\hat{f}} = \frac{1}{N}A^*(\omega)\vec{f}$, $\vec{\hat{f}}$ is the vector of Fourier coefficients. In addition, $\vec{f} = A(\omega)\vec{\hat{f}}$. Thus,

$$D\vec{u} = \vec{f} \Rightarrow DA(\omega)\vec{\hat{u}} = A(\omega)\vec{\hat{f}} \Rightarrow \frac{1}{N}A(\omega)^* DA(\omega)\vec{\hat{u}} = \vec{\hat{f}}$$

Moreover $\frac{1}{N}A^*(\omega)DA(\omega)$ can be simplified to a diagonal matrix. In the following, we let $\omega^{(m)} = (\omega^{0\times m}, \omega^{1\times m}, ..., \omega^{(N-1)\times m})^T$ for $m = 0, 1, ..., N-1$

3.1 Eigen-decomposition of circulant matrix

Proposition 3.1. Suppose C is a circulant $N \times N$ matrix and its first row is $(c_0, c_{N-1}, ..., c_1)^T$, then $\omega^{(m)}$ is an eigenvector of C corresponding to the eigenvalue $\lambda_m = N\text{DFT}(\vec{c})(m)$, where $\vec{c} = (c_0, c_1, ..., c_{N-1})^T$

So each column of $A(\omega)$ is an eigenvector of D. What about $A^*(\omega)$? note that

$$
\begin{aligned}
A^*(\omega) &= \begin{bmatrix} \omega^{0*0} & \omega^{0*(-1)} & \cdots & \omega^{0*[-(N-1)]} \\ \omega^{1*0} & \omega^{1*(-1)} & \cdots & \omega^{1*[-(N-1)]} \\ \vdots & \vdots & \ddots & \vdots \\ \omega^{(N-1)*0} & \omega^{(N-1)*(-1)} & \cdots & \omega^{(N-1)*[-(N-1)]} \end{bmatrix} \\
&= \begin{bmatrix} \omega^{0*0} & \omega^{0*(N-1)} & \cdots & \omega^{0*1} \\ \omega^{1*0} & \omega^{1*(N-1)} & \cdots & \omega^{1*1} \\ \vdots & \vdots & \ddots & \vdots \\ \omega^{(N-1)*0} & \omega^{(N-1)*(N-1)} & \cdots & \omega^{(N-1)*1} \end{bmatrix} \\
&= \left(\begin{bmatrix} | & | & \cdots & | \\ \omega^{(0)} & \omega^{(N-1)} & \cdots & \omega^{(1)} \\ | & | & \cdots & | \end{bmatrix} \right)
\end{aligned}
$$

In all,

$$
\begin{aligned}
\frac{1}{N} A^*(\omega) D A(\omega) \vec{\hat{u}} &= \frac{1}{N} A^*(\omega) A(\omega) diag(\lambda_0, \lambda_1, \lambda_2, ..., \lambda_{N-1}) \vec{\hat{u}} \\
&= diag(\lambda_0, \lambda_1, \lambda_2, ..., \lambda_{N-1}) \vec{\hat{u}} \\
&= (\lambda_0 \cdot \hat{u}(0), \quad \lambda_1 \cdot \hat{u}(1), \cdots, \quad \lambda_{N-1} \cdot \hat{u}(N-1))^T \\
&= (\hat{f}(0), \hat{f}(1), \cdots, \hat{f}(N-1))^T
\end{aligned}
$$

After obtaining $\{\hat{u}(k)\}$, we can recover $\{u(x_j\}$ by $u(x_j) = \sum_{k=0}^{N-1} \hat{u}(k) e^{ik\frac{2\pi}{N}j}$. This is a general framework for transforming a circulant linear system in the spatial domain to a diagonal linear system in the frequency domain. You are not suggested to recite the conclusion above. Instead, you should try to understand the spirit hidden behind it.

The two essential things of applying this method are

1. our goal is to obtain $\{\hat{u}(k)\}$, and each $\hat{u}(k)$ is a constant multiple of $\hat{f}(k)$, this constant only depends on the circulant matrix, or equivalently, the finite difference scheme you design.

2. These constant factors are eigenvalues of the circulant matrix and you can easily compute the eigenvalues by computing the first row of matrix-vector product.

Before further discussing the potential confusion induced, we would like to illustrate this process with a concrete example in case some students might find it difficult to put this abstract process into practice.

Question 3.1. Given a partial differential equation $(\frac{d^2}{dx^2} + 2\frac{d}{dx})u = f$ with observations $f(x_j)$ and $x_j = \frac{2\pi j}{N}, j = 0, \cdots, N-1$. Suppose we approximate $u''(x_j)$ by $\frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{h^2}$ and $u'(x_j)$ by $\frac{u(x_{j+1}) - u(x_{j-1})}{2h}$, Could you obtain the solution u, and is there only one solution? If not, why?

From this example, you should have a taste of solving the ODE using a finite difference scheme. The idea is that

$$\begin{pmatrix} u(x_0) \\ u(x_1) \\ \vdots \\ u(x_{N-1}) \end{pmatrix} = \hat{u}(0) \cdot \begin{pmatrix} \omega^{0\times 0} \\ \omega^{0\times 1} \\ \vdots \\ \omega^{0\times (N-1)} \end{pmatrix} + \hat{u}(1) \cdot \begin{pmatrix} \omega^{1\times 0} \\ \omega^{1\times 1} \\ \vdots \\ \omega^{1\times (N-1)} \end{pmatrix} + \cdots + \hat{u}(N-1) \cdot \begin{pmatrix} \omega^{N-1\times 0} \\ \omega^{(N-1)\times 1} \\ \vdots \\ \omega^{(N-1)\times (N-1)} \end{pmatrix}$$
$$= \hat{u}(0) \cdot \omega^{(0)} + \hat{u}(1) \cdot \omega^{(1)} + \cdots + \hat{u}(N-1) \cdot \omega^{(N-1)}$$

And

$$D \begin{pmatrix} u(x_0) \\ u(x_1) \\ \vdots \\ u(x_{N-1}) \end{pmatrix} = \hat{u}(0) \cdot D\omega^{(0)} + \hat{u}(1) \cdot D\omega^{(1)} + \cdots + \hat{u}(N-1) \cdot D\omega^{(N-1)}$$
$$= \hat{u}(0) \cdot d_0 \cdot \omega^{(0)} + \hat{u}(1) \cdot d_1 \cdot \omega^{(1)} + \cdots + \hat{u}(N-1) \cdot d_{N-1} \cdot \omega^{(N-1)}$$

So if the eigenvalue d_j of D is nonzero, $\hat{u}(j) = \frac{\hat{f}(j)}{d_j}$. However, what if some of the d_j are zero? and Why some of the d_j are zero? We will discuss it in the tutorial.