## **LECTURE 19: DERIVATIVE APPROXIMATION**

Having introduced approximations of elements of vector spaces, a natural question is the approximation of transformations of such objects or operator approximation. An operator is understood here as a mapping from a domain vector space  $\mathcal{U} = (U, S, +, \cdot)$  to a co-domain vector space  $\mathcal{V} = (V, S, +, \cdot)$ , and the operator  $\mathcal{L}: U \to V$  is said to be linear if for any scalars  $c_1, c_2 \in S$  and vectors  $u_1, u_2 \in U$ ,

$$\mathscr{L}(c_1u_1 + c_2u_2) = c_1 \mathscr{L}(u_1) + c_2 \mathscr{L}(u_2)$$

i.e., the image of a linear combination is the linear combination of the images. Linear algebra considers the case of finite dimensional vector spaces, such as  $U = \mathbb{R}^m$ ,  $V = \mathbb{R}^n$ , in which case a linear operator is represented by a matrix  $L \in \mathbb{R}^{m \times n}$ , and satisfies

$$\boldsymbol{L}(c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2) = c_1\boldsymbol{L}\boldsymbol{u}_1 + c_2\boldsymbol{L}\boldsymbol{u}_2.$$

In contrast, the focus here is on infinite-dimensional function spaces such as  $C^r(\mathbb{R})$  (cf. Tab. 1, L18), the space of functions with continuous derivatives up to order *r*. Common linear operator examples include:

**Differentiation.**  $\mathscr{L}f = \partial^k f / \partial t^k$ ,  $\mathscr{L}: C^r(\mathbb{R}) \to C^{r-k}(\mathbb{R})$ .

**Riemann integration.**  $\mathscr{L}f = \int_{a}^{b} \omega(t) f(t) dt$ ,  $\mathscr{L}: C(\mathbb{R} \setminus \Delta) \to \mathbb{R}$ , where  $\Delta$  is a set of measure zero.

**Linear differential equation.**  $\mathscr{L}y = \sum_{i=0}^{k} a_i(t) y^{(i)} = f(t), \ \mathscr{L}: C^r(\mathbb{R}) \to C^{r-k}(\mathbb{R}).$ 

## 1. Numerical differentiation based upon polynomial interpolation

A general approach to operator approximation is to simply introduce an approximation of the function the operator acts upon,  $f \cong p$ ,

$$\mathscr{L}f \cong \mathscr{L}p.$$

**Monomial basis.** As an example consider the polynomial interpolant of *f* based upon data  $\mathcal{D} = \{(x_i, y_i = f(x_i)), i = 0, ..., n\},\$ 

$$p(t) = \begin{bmatrix} 1 & t & t^2 & \dots & t^n \end{bmatrix} \boldsymbol{c},$$

with coeffcients c determined as the solution of the interpolation conditions

$$Mc = y$$
,

with notations

$$\boldsymbol{M} = \begin{bmatrix} \mathbf{1} \ \boldsymbol{x} \ \boldsymbol{x}^2 \ \dots \ \boldsymbol{x}^n \end{bmatrix}, \boldsymbol{x}^k = \begin{bmatrix} x_0^k \ \dots \ x_n^k \end{bmatrix}^T, \boldsymbol{y} = \begin{bmatrix} y_0 \ \dots \ y_n \end{bmatrix}^T.$$

Differentiation of  $f(\mathcal{L} = d/dt)$  can be approximated as

$$\frac{\mathrm{d}}{\mathrm{d}t}f \cong \frac{\mathrm{d}}{\mathrm{d}t}p = \begin{bmatrix} 0 & 1 & 2t & \dots & nt^{n-1} \end{bmatrix} c.$$

It is often of interest to express the result of applying an operator directly in terms of known information on f. Formally, in the case of differentiation,

$$\frac{\mathrm{d}}{\mathrm{d}t}f \cong \begin{bmatrix} 0 & 1 & 2t & \dots & nt^{n-1} \end{bmatrix} \boldsymbol{M}^{-1}\boldsymbol{y},$$

allowing the identification of a differentiation approximation operator  $\mathcal{D}$ 

$$\frac{\mathrm{d}}{\mathrm{d}t}f\cong \mathcal{D}(\mathbf{y}), \mathcal{D}=\begin{bmatrix}0 \ 1 \ 2t \ \dots \ nt^{n-1}\end{bmatrix}\boldsymbol{M}^{-1}.$$

This formulation explicitly includes the inversion of the sampled basis matrix M, and is hence not computationally efficient. Alternative formulations can be constructed that carry out some of the steps in computing  $M^{-1}$  analytically.

Newton basis (finite difference calculus). An especially useful formulation for numerical differentiation arises from the Newton interpolant of data  $\mathcal{D} = \{(x_i = ih, y_i = f(x_i)), i = 0, ..., n\}, f : \mathbb{R} \to \mathbb{R}, f \in C^{(n+1)}(\mathbb{R}), f \in C^{(n+1)}(\mathbb{R})\}$ 

 $f(t) \cong p(t) = [y_0] + [y_1, y_0](t - x_0) + \dots + [y_n, y_{n-1}, \dots, y_0](t - x_0) \cdot (t - x_1) \cdot \dots \cdot (t - x_{n-1}).$ 

For equidistant sample points  $x_i = ih$ , the Newton interpolant can be expressed as an operator acting upon the data. Introduce the translation operator

$$Ef(t) = f(t+h).$$

Repeated application of the translation operator leads to

$$E^k f(t) = E(E^{k-1}f(t)) = \cdots = f(t+kh),$$

and the identity operator is given by

$$If(t) = f(t) = E^0 f(t) \Rightarrow I = E^0.$$

Finite differences of the function values are expressed through the forward, backward and central operators

$$\Delta = E - I, \nabla = I - E, \delta = E^{1/2} - E^{-1/2},$$

leading to the formulas

$$\Delta f(t)=f(t+h)-f(t), \nabla f(t)=f(t)-f(t-h), \delta f(t)=f(t+h/2)-f(t-h/2)$$

Applying the above to the data set  $\mathcal{D}$  leads to

$$\Delta y_i = y_{i+1} - y_i, \nabla y_i = y_i - y_{i-1}, \, \delta y_i = y_{i+1/2} - y_{i-1/2}.$$

The divided differences arising in the Newton can be expressed in terms of finite difference operators,

$$[y_1, y_0] = \frac{y_1 - y_0}{h} = \frac{1}{h} \Delta y_0, [y_2, y_1, y_0] = \frac{[y_2, y_1] - [y_1, y_0]}{2h} = \frac{\Delta y_1 - \Delta y_0}{2h^2} = \frac{\Delta^2 y_0}{2h^2},$$

or in general

$$[y_k,\ldots,y_1,y_0]=\frac{\Delta^k}{k!h^k}y_0.$$

Using the above and rescaling the variable t in the Newton basis  $\mathcal{N} = \{1, t-x_0, (t-x_0), (t-x_1), \ldots\}$  in units of the step size  $t = \alpha h + x_0$  leads to

$$p(t(\alpha)) = P(\alpha) = \left(I + \alpha \frac{\Delta}{1!} + \alpha (\alpha - 1) \frac{\Delta^2}{2!} + \dots + \alpha (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n) \frac{\Delta^n}{n!}\right) y_0.$$
(1)

The generalized binomial series states

$$(1+x)^{\alpha} = \sum_{k=0}^{\infty} {\binom{\alpha}{k} x^{k}},$$

$$\begin{pmatrix} \alpha \\ \ddots \end{pmatrix} = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{\alpha-k-1}$$
(2)

with

$$\begin{pmatrix} \alpha \\ k \end{pmatrix} = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}$$

the generalized binomial coefficient. The operator acting upon  $y_0$  in (1) can be interpreted as the truncation at order n

$$P(\alpha) \cong (I + \Delta)^{\alpha} y_0 = \mathscr{F}_{\alpha} y_0,$$

of the operator  $(I + \Delta)^{\alpha}$  defined through (2) by the substitutions  $1 \rightarrow I, x \rightarrow \Delta$ . The operator  $\mathscr{F}_{\alpha} = (I + \Delta)^{\alpha}$  can be interpreted as the interpolation operator with equidistant sampling points, with  $P(\alpha)$  its truncation to order n. Reversing the order of the sampling points leads to the Newton interpolant

$$p(t) = [y_n] + [y_{n-1}, y_n](t - x_n) + \dots + [y_0, y_1, \dots, y_n](t - x_n)(t - x_{n-1}) \cdot \dots \cdot (t - x_1)$$

The divided differences can be expressed in terms of the backward operator as

$$[y_{n-1}, y_n] = \frac{y_{n-1} - y_n}{h} = -\frac{1}{h} \nabla y_n, [y_{n-2}, y_{n-1}, y_n] = \frac{[y_{n-2}, y_{n-1}] - [y_{n-1}, y_n]}{2h} = -\frac{\nabla y_{n-1} - \nabla y_n}{2h^2} = \frac{\nabla^2 y_n}{2h^2},$$

leading to an analogous expression of the interpolation operator in terms backward finite differences

$$p(t(\alpha)) = P(\alpha) = \left(I - \alpha \frac{\nabla}{1!h} + \alpha(\alpha - 1)\frac{\nabla^2}{2!h^2} + \dots + (-1)^n \alpha(\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n}\right) y_n \cong (I - \nabla)^{\alpha} y_n = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1) \cdot \dots \cdot (\alpha - 1 + n)\frac{\nabla^n}{n!h^n} = \mathcal{B}_{\alpha} y_n \cdot (\alpha - 1) \cdot \dots \cdot ($$

Differentiation of the interpolation expressed in terms of forward finite differences gives

$$f'(t) \cong \frac{\mathrm{d}\alpha}{\mathrm{d}t} P(\alpha) = \frac{\mathrm{d}\alpha}{\mathrm{d}t} P'(\alpha) \cong \frac{1}{h} \frac{\mathrm{d}}{\mathrm{d}\alpha} \mathscr{F}_{\alpha} y_0 = \frac{1}{h} \left[ \ln(I + \Delta) \right] (I + \Delta)^a y_0 \cong \frac{1}{h} \ln(I + \Delta) P(\alpha).$$

The particular interpolant  $P(\alpha)$  is irrelevant, leading to the operator identity

$$\frac{\mathrm{d}}{\mathrm{d}t} \cong \frac{1}{h} \ln(I + \Delta).$$

For |x| < 1, the power series expansions are

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln(1+x) = \frac{1}{1+x} = 1 - x + x^2 - \dots \Rightarrow \ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots + (-1)^{k+1}\frac{x^k}{k} + \dots,$$

are uniformly convergent, leading to the expression

$$\frac{\mathrm{d}}{\mathrm{d}t} \cong \frac{1}{h} \left( \Delta - \frac{1}{2} \Delta^2 + \frac{1}{3} \Delta^3 - \ldots + (-1)^k \frac{1}{k} \Delta^k + \cdots \right),$$

stating that the (continuum) differentiation operator can be approximated by an infinite series of finite difference operations, recovered exactly in the  $h \rightarrow 0$  limit. Denote by  $D_k^+$  the truncation at term k of the above operator series such that

$$f'(x_0) \cong D_k^+(f)(x_0) = \frac{1}{h} \left( \Delta - \frac{1}{2} \Delta^2 + \frac{1}{3} \Delta^3 - \dots + (-1)^k \frac{1}{k} \Delta^k \right) y_0.$$

Truncation at k = 1, 2, 3 leads to the expressions

$$D_{1}^{+}(f) = \frac{f(h+t) - f(t)}{h}, D_{2}^{+}(f) = \frac{4f(h+t) - f(2h+t) - 3f(t)}{2h}, D_{3}^{+}(f) = \frac{18f(h+t) - 9f(2h+t) + 2f(3h+t) - 11f(t)}{6h}.$$

The  $h \rightarrow 0$  limit of divided differences is given by

$$\lim_{h \to 0} [y_k, y_{k-1}, \dots, y_0] = \lim_{h \to 0} \left( \frac{1}{k! h^k} \Delta^k y_0 \right) = \frac{1}{k!} f^{(k)}(x_0),$$

such that for small finite h > 0,

$$\Delta^k y_0 \cong h^k f^{(k)}(x_0).$$

The resulting derivative approximation error is of order k,

$$e_k^+(t) = D_k^+(f)(t) - f'(t) = \frac{(-1)^{k+1}h^k}{k+1}f^{(k+1)}(t) = \mathcal{O}(h^k).$$

The analogous expression for backward differences is

$$\frac{\mathrm{d}}{\mathrm{d}t} \cong -\frac{1}{h}\ln(I-\nabla) = \frac{1}{h}\left(\nabla + \frac{1}{2}\nabla^2 + \frac{1}{3}\nabla^3 + \ldots + \frac{1}{k}\nabla^k + \cdots\right),$$

and the first few truncations are

$$D_{1}^{-}(f) = \frac{f(t-h) - f(t)}{h}, D_{2}^{-}(f) = \frac{-f(t-2h) + 4f(t-h) - 3f(t)}{2h}, D_{3}^{-}(f) = \frac{2f(t-3h) - 9f(t-2h) + 18f(t-h) - 11f(t)}{6h}$$
with errors

$$e_{k}^{-}(t) = D_{k}^{-}(f)(t) - f'(t) = \frac{h^{k}}{k}f^{(k+1)}(t) = \mathcal{O}(h^{k}).$$

The above operator identities can be inverted to obtain

$$\Delta = E - I = \exp\left(h\frac{\mathrm{d}}{\mathrm{d}t}\right) - I, \nabla = I - E^{-1} = I - \exp\left(-h\frac{\mathrm{d}}{\mathrm{d}t}\right),$$

leading to

$$E = \exp\left(h\frac{\mathrm{d}}{\mathrm{d}t}\right) = 1 + h\frac{\mathrm{d}}{\mathrm{d}t} + \frac{1}{2}\left(h\frac{\mathrm{d}}{\mathrm{d}t}\right)^2 + \dots + \frac{1}{k!}\left(h\frac{\mathrm{d}}{\mathrm{d}t}\right)^k + \dots + \frac{1}{k!}\left(h\frac{\mathrm{d$$

this time expressing the finite translation operator as an infinite series of continuum differentiation operations. This allows expressing the central difference operator as

$$\delta = E^{1/2} - E^{-1/2} = \exp\left(\frac{h}{2}\frac{d}{dt}\right) - \exp\left(-\frac{h}{2}\frac{d}{dt}\right) = 2\sinh\left(\frac{h}{2}\frac{d}{dt}\right)$$

and approximations of the derivative based on centered differencing are obtained from

$$\frac{d}{dt} \approx \frac{2}{h} \operatorname{arcsinh}\left(\frac{\delta}{2}\right) = \frac{1}{h} \left(\delta - \frac{\delta^3}{24} + \frac{3\delta^5}{640} - \frac{5\delta^7}{7168} + \frac{35\delta^9}{294912} - \cdots\right)$$

An advantage of the centered finite differences (surmised from the odd power series) is a higher order of accuracy

$$_{k} = D_{k}f(f) - f'(t) = \mathcal{O}(h^{2k})$$

е Higher order derivative are obtained by repeated application of the operator series, e.g.,

$$\frac{d^2}{dt^2} = \frac{d}{dt} \cdot \frac{d}{dt} = \frac{1}{h^2} \left( \Delta - \frac{1}{2} \Delta^2 + \frac{1}{3} \Delta^3 - \cdots \right)^2 = \frac{1}{h^2} \left( \Delta^2 - \Delta^3 + \frac{11}{12} \Delta^4 - \cdots \right)^2.$$

## 2. Taylor series methods

An alternative derivation of the above finite difference formulas is to construct a linear combination of function values

$$L_{m}^{n}f(t) = \sum_{k=-m}^{n} c_{k}f(t+kh) = \left(\sum_{k=-m}^{n} c_{k}E^{k}\right)f(t),$$

and determine the coefficients  $c_k$  such that the  $p^{\text{th}}$  derivative is approximated to order q

$$f^{(p)}(t) = L_m^n f(t) + \mathcal{O}(h^q).$$

For example, for m = 0, n = 1, carrying out Taylor series expansions gives

$$f(t+h) = f(t) + hf'(t) + \frac{1}{2}h^2 f''(t) + \cdots$$
  
$$f(t) = f(t).$$

Eliminating f(t) by multiplying the first equation by  $c_1 = 1$  and the second by  $c_0 = -1$  recovers the forward finite difference formula

$$f'(t) = \frac{f(t+h) - f(t)}{h} + \mathcal{O}(h).$$

## 3. Numerical differentiation based upon piecewise polynomial interpolation

**B-spline basis.** The above example used a truncation of the monomial basis  $\mathcal{M}_n(t) = \{1, t, \dots, t^n\}$ . Analogous results are obtained when using a different basis. Consider the equidistant sample points  $x_i = ih + x_0$ , data  $\mathcal{D} = \{(x_i, y_i = f(x_i), y_i = f$  $i = 0, 1, \dots, n$  and the first-degree *B*-spline basis

$$\mathscr{B}_{n,1}(t) = \{B_{0,1}(t), B_{1,1}(t), \dots, B_{n,1}(t)\},\$$

in which case the linear piecewise interpolant is expressed as

$$p(t) = \sum_{i=0}^{n} y_i B_{i,1}(t),$$

and over interval  $[x_{i-1}, x_i]$  reduces to

$$p_i(t) = y_{i-1}B_{i-1,1}(t) + y_iB_{i,1}(t) = y_{i-1} \cdot \frac{x_i - t}{x_i - x_{i-1}} + y_i \cdot \frac{t - x_{i-1}}{x_i - x_{i-1}}$$

Differentiation recovers the familiar slope expression

$$p_i'(t) = \frac{y_i - y_{i-1}}{x_i - x_{i-1}} = \frac{y_i - y_{i-1}}{3h}$$

At the nodes, a piecewise linear spline is discontinuous, hence the derivative is not defined, though one could consider the one-sided limits. Evaluation of derivatives at midpoints  $t_i = (x_{i-1} + x_i)/2 = (i-1)h + h/2 + x_0$ , i = 1, 2, ..., n, leads to

with  $\boldsymbol{D} \in \mathbb{R}^{n \times (n+1)}$ .