The physical world is continuous, yet numerical analysis frequently requires that continuous functions be approximately calculated on a finite set of points. The principal tool for solving differential and integral equations numerically is to discretize the underlying equations on a lattice, and to find the solution on that lattice. The spacing between lattice points and the total size of the lattice are chosen so that the discrete solution provides an accurate representation of the continuous one.

For reconstructing probability distributions from a finite data sample, the most natural approach is to allow the data points themselves to define the lattice. Consequently, we have to compute integration and differentiation operators on irregular lattices. That task is the subject of this note.

1 A one-dimensional example

To motivate our approach, consider first an infinite, evenly-spaced lattice of points in one dimension. A function $f(x)$ on continuous space has value f_n at the nth lattice point x_n. Adjacent lattice points are separated by a distance a. Then, conventional practice would tell us that the integral of f may be approximated by

$$\int \mathrm{d}x\, f(x) = \sum_n a f_n + O(a^3), \tag{1}$$

while the derivative and laplacian are

$$f'\left(\frac{x_n + x_{n+1}}{2}\right) = \frac{f_{n+1} - f_n}{a} + O(a^3) \tag{2}$$

and

$$f''(x_n) = \frac{\frac{f_{n+1}-f_n}{a} - \frac{f_n - f_{n-1}}{a}}{a} + O(a^4) = \frac{f_{n+1} - 2f_n + f_{n-1}}{a^2} + O(a^4), \tag{3}$$

respectively. (The derivative is best approximated at intermediate points, where the error due to discretization is of order $O(a^3)$ rather than $O(a^2)$.)

Eqs. (1) and (2) are *exact* if f happens to be piecewise linear, with changes in slope only at the points of the regular lattice. This fact suggests a simple method for generalizing these equations to an arbitrary lattice: to calculate the exact integrals and derivatives for piecewise linear functions on the lattice. This results in the expressions

$$\int \mathrm{d}x\, f(x) \approx \sum_n (x_{n+1} - x_n)\frac{f_n + f_{n+1}}{2} = \sum_n \frac{x_{n+1} - x_{n-1}}{2} f_n, \tag{4}$$

and

$$f'\left(\frac{x_n + x_{n+1}}{2}\right) \approx \frac{f_{n+1} - f_n}{x_{n+1} - x_n}. \tag{5}$$

To define the second derivative, we might simply iterate Eq. (5):

$$f''(x_n) \approx \frac{\frac{f_{n+1}-f_n}{x_{n+1}-x_n} - \frac{f_n - f_{n-1}}{x_n - x_{n-1}}}{\frac{x_{n+1}+x_n}{2} - \frac{x_n + x_{n-1}}{2}}. \tag{6}$$

In other words,

$$\mathrm{d}x\, f''(x_n) \approx \frac{f_{n+1} - f_n}{x_{n+1} - x_n} - \frac{f_n - f_{n-1}}{x_n - x_{n-1}}, \tag{7}$$

where the identification $\mathrm{d}x \approx \frac{x_{n+1}-x_{n-1}}{2}$ is made from Eq. (4). This simple procedure for defining the laplacian encounters some difficulty in higher dimensions, where the divergence operator differs from the gradient. A more general strategy will be to define the laplacian using integration-by-parts, in terms of the integration and differentiation operators we've already defined:

$$\int \mathrm{d}x\, f'^2 = -\int \mathrm{d}x\, f'' f, \tag{8}$$

and

$$\sum_n \frac{(f_{n+1} - f_n)^2}{x_{n+1} - x_n} = -\sum_n \left[\frac{f_{n+1} - f_n}{x_{n+1} - x_n} - \frac{f_n - f_{n-1}}{x_n - x_{n-1}}\right] f_n, \tag{9}$$

which yields the same result as Eq. (7).

2 Differentiation and integration on higher-dimensional lattices

In higher dimensions, the same basic procedure can be used: define the integration, gradient, and laplacian operators by calculating them for functions that are piecewise linear on the lattice. The only difficulty lies in defining "piecewise linear" on an arbitrary lattice. Fortunately, there exists a construction which allows this notion to be defined in a natural way. For an arbitrary set of points $\mathbf{x}_n$, the *Delaunay triangulation* consists of a complete, non-overlapping set of polyhedra whose vertices are taken from the $\mathbf{x}_n$, and which contain none of the $\mathbf{x}_n$ in their interiors. From a set of function values f_n known only at the lattice points $\mathbf{x}_n$, one may define a corresponding function $\hat{f}$ which is continuous everywhere, linear on every polyhedron of the Delaunay triangulation, and which changes slope only at the faces (edges, in two dimensions) of the polyhedra.

Consider a d-dimensional polyhedron whose vertices are specified by $\mathbf{x}_1 \dots \mathbf{x}_{d+1}$. The points $(\mathbf{x}_i, f_i)$ lie on a hyperplane in $d+1$ dimensions which may be parametrized at arbitrary points $(\mathbf{x}, \hat{f})$ as

$$(\mathbf{x}, \hat{f}) \cdot \hat{\mathbf{n}} = c, \tag{10}$$

for some constant c and the unit normal vector $\hat{\mathbf{n}}$ in $d+1$ dimensions. This expression may be used to calculate $\hat{f}(\mathbf{x})$ only when $\mathbf{x}$ is in the interior of the given polyhedron. Our task, then, is to determine the constants c and $\hat{\mathbf{n}}$ from the known values $(\mathbf{x}_i, f_i)$. First, we note that Eq. (10) may be re-written as

$$(\mathbf{x}, \hat{f}) \cdot \mathbf{n} = 1, \tag{11}$$

where $\mathbf{n} = \hat{\mathbf{n}}/c$ is now some $d+1$-vector with arbitrary normalization. Let $\mathbf{M}$ be the $d+1 \times d+1$ matrix with ith row $(\mathbf{x}_i, f_i)$; then Eq. (11), when applied to the vertices of the polyhedron, is simply

$$\mathbf{M}\mathbf{n} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{12}$$

$\mathbf{n}$ may therefore be calculated by solving this linear equation. Once $\mathbf{n}$ is known for each polyhedron, the value of $\hat{f}$ is determined at all points.

The integration and gradient operators may now be calculated for each polyhedron.

Lemma 1 *The volume of the polyhedron specified by the points* $\mathbf{x}_1, \dots, \mathbf{x}_{d+1}$ *is* $1/d! |det \mathbf{M}_{d+1}|$, *where* $\mathbf{M}_{d+1}$ *is the matrix* $\mathbf{M}$ *with column* $d+1$ *replaced with ones.*

Lemma 2 *The center-of-mass of a polyhedron specified by the points* $\mathbf{x}_1, \dots, \mathbf{x}_{d+1}$ *is the mean of the* $\mathbf{x}_i$.

Lemma 3 *The integral of a linear function* $\hat{f}$ *over a polyhedron specified by the points* $\mathbf{x}_1, \dots, \mathbf{x}_{d+1}$ *is the volume of the polyhedron times the mean of the values* f_i *at the vertices of the polyhedron.*

The integral may be expressed as the mean value of $\hat{f}$ over the polyhedron times the volume of the polyhedron. Since $\hat{f}$ is linear, its mean value is equal to its value at the center-of-mass of the polyhedron. But the point $(\frac{1}{d+1}\sum \mathbf{x}_i, \frac{1}{d+1}\sum f_i)$ is on the hyperplane determined by $\hat{f}$, so the value at the center-of-mass is simply $\frac{1}{d+1}\sum f_i$.

Since $\hat{f}$ is linear over the polyhedron, its gradient is constant, and from Eq. (11) is given by

$$\partial_i \hat{f} = -\frac{n_i}{n_{d+1}}, \tag{13}$$

where $\mathbf{n}$ is determined by Eq. (12). It follows that the kinetic energy can be expressed as

$$\int_P \mathrm{d}\mathbf{x}\, (\nabla \hat{f})^2 = \frac{1}{d! n_{d+1}^2} |det \mathbf{M}_{d+1}| \sum_{i=1}^{d} n_i^2. \tag{14}$$

The result can be used to define the laplacian operator, as in Eq. (9). However, a bit more work is required to convert these results into operator expressions.

3 Operators for differentiation and integration

The results of the previous section suffice to compute integrals and derivatives of a given function. Suppose, however, that our goal is to solve some linear integral or partial differential equation on the lattice. Then, the function itself is not known; instead, the function is determined by the solution of an equation $\mathcal{O}\mathbf{f} = \mathbf{j}$, where $\mathcal{O}$ is an operator which, when applied to $\mathbf{f}$, computes the appropriate derivatives or integrals. Our goal therefore is to extract such operators from the results of the previous section.

The integration operator is extremely simple. For a lattice determined by the points $\mathbf{x}_1, \ldots, \mathbf{x}_N$, the total integral can be represented as

$$\int \mathrm{d}\mathbf{x}\, \hat{f} = \sum_i \mathrm{vol}(P_i)(\text{mean of the } f_j \text{ at the vertices of } P_i). \tag{15}$$

This may be rearranged as

$$\int \mathrm{d}\mathbf{x}\, \hat{f} = \mathbf{v} \cdot \mathbf{f}, \tag{16}$$

where

$$v_i = \frac{1}{d+1} \sum_{P_j \text{ with vertex } \mathbf{x}_i} \mathrm{vol}(P_j). \tag{17}$$

The gradient operator requires a bit more effort. Each polyhedron will have an associated $d \times d+1$ gradient matrix $\mathbf{G}$. Suppose the vertices of a given polyhedron are numbered $1, \ldots, d+1$. Eqs. (13) and (12) can be used in conjunction with Cramer's rule to say that

$$\partial_i \hat{f} = -\frac{\det \mathbf{M}_i}{\det \mathbf{M}_{d+1}}, \tag{18}$$

where $\mathbf{M}_j$ is the matrix $\mathbf{M}$ with the jth column replaced by 1s. (Note that $\det \mathbf{M}_{d+1}$ is the determinant which appears in the volume of the polyhedron.) This expression is linear in the f_j. The coefficient of f_j can be extracted, yielding

$$\partial_i \hat{f} = \sum_j \mathbf{G}_{ij} f_j, \tag{19}$$

where

$$\mathbf{G}_{ij} = \frac{(-1)^{j+d} \det \mathbf{M}_i(j|d+1)}{\det \mathbf{M}_{d+1}}, \tag{20}$$

and $\mathbf{A}(i|j)$ is the matrix $\mathbf{A}$ with the ith row and jth column deleted.

With the gradient in hand, the kinetic energy on a given polyhedron is

$$\int_P \mathrm{d}\mathbf{x}\, (\nabla \hat{f})^2 = \mathbf{f}_P^T \mathbf{T}_P \mathbf{f}_P, \tag{21}$$

where $\mathbf{T}_P = \mathrm{vol}(P)\mathbf{G}^T\mathbf{G}$. Explicitly,

$$(\mathbf{T}_P)_{ij} = \frac{(-1)^{i+j}}{d!|\det \mathbf{M}_{d+1}|} \sum_k \det \mathbf{M}_k(i|d+1) \det \mathbf{M}_k(j|d+1). \tag{22}$$

The kinetic energy operators $\mathbf{T}_P$ for each polyhedron may be combined into a single operator $\mathbf{T}$ for the entire lattice: $\mathbf{T}_{ij}$ is the sum of $(\mathbf{T}_P)_{ij}$ over all polyhedra containing the given pair of vertices.

The kinetic energy operator $\mathbf{T}$ can be used to define the Laplacian $\mathbf{L}$, using integration-by-parts as in Eq. (9):

$$\mathbf{f}^T\mathbf{T}\mathbf{f} = \int \mathrm{d}\mathbf{x}\, (\nabla \hat{f})^2 = -\int \mathrm{d}\mathbf{x}\, \hat{f} \nabla^2 \hat{f} \equiv -\mathbf{f}^T \mathbf{V}\mathbf{L}\mathbf{f}, \tag{23}$$

where $\mathbf{V}$ is the diagonal matrix with the values of the volume operator $\mathbf{v}$ along the diagonal. Consequently, we might set $\mathbf{L} = -\mathbf{V}^{-1}\mathbf{T}$. This apparently sensible choice suffers from one defect: $\mathbf{L}$ is not guaranteed to be symmetric unless $\mathbf{V}$ and $\mathbf{T}$ commute. It turns out that $\mathbf{V}$ and $\mathbf{T}$ do indeed commute on a lattice with no boundary (either infinite in extent, or with periodic boundary conditions) how to prove this?. If the lattice does have points which lie on the boundary of the Delaunay triangulation, then in any event an additional term must be inserted into the integration-by-parts, and definition of the Laplacian by this device becomes more difficult. It should be pointed out than an alternative definition $L_{ij} = -\frac{2T_{ij}}{v_i+v_j}$ is guaranteed to be symmetric, and satisfies Eq. (23). Away from boundaries, this definition is equivalent to the choice $\mathbf{L} = -\mathbf{V}^{-1}\mathbf{T}$ proof?. However, this definition suffers from other problems, as can be illustrated by a simple example: consider solving the Helmholtz equation

$$(-\nabla^2 + k^2)f = j \tag{24}$$

on an arbitrary lattice. This equation can be derived as the equation of motion for the minimum of the action

$$S[f] = \int d\mathbf{x} \left[\frac{1}{2}(\nabla f)^2 + k^2 f^2 - fj\right]. \tag{25}$$

The discrete version of the action on the lattice is

$$S(\mathbf{f}) = \mathbf{f}^T \left[\left(\frac{1}{2}\mathbf{T} + k^2\mathbf{V}\right)\mathbf{f} - \mathbf{V}\mathbf{j}\right]. \tag{26}$$

Consequently, the discrete Helmholtz equation is

$$(\mathbf{T} + k^2\mathbf{V})\mathbf{f} = \mathbf{V}\mathbf{j}. \tag{27}$$

This agrees with the equation $(-\mathbf{L}+k^2)\mathbf{f} = \mathbf{j}$ that might have been suspected naively, given the identification $\mathbf{L} = -\mathbf{V}^{-1}\mathbf{T}$. The alternative choice $L_{ij} = -\frac{2T_{ij}}{v_i+v_j}$ proves to be incorrect, unless the two happen to be equivalent which happens when there is no boundary?.

The kinetic energy operator $\mathbf{T}$ is therefore the fundamental second-derivative operator. An operator corresponding to a fourth derivative may be easily constructed; the proper choice is

$$\int d\mathbf{x}\,(\nabla^2 f)^2 = \mathbf{f}^T\mathbf{T}\mathbf{V}^{-1}\mathbf{T}\mathbf{f}. \tag{28}$$