Massachusetts Institute of Technology

Functionals and the variational derivative

Using energies to make sense of PDEs

18.303 Linear Partial Differential Equations: Analysis and Numerics

Functionals

Assume we have a function $V \ni f : \mathbb{R}^N \to \mathbb{R}$ that takes N-dimensional vectors and maps them to real numbers.

Functionals

Assume we have a function $V \ni f : \mathbb{R}^N \to \mathbb{R}$ that takes N-dimensional vectors and maps them to real numbers.

Functionals are a generalization of this concept. Formally we have $F : V \to \mathbb{R}$. We can write these using an integral as

$$F[f] = \int_{\Omega} g(f(\mathbf{x}), \dots) \mathrm{d}\mathbf{x}.$$

Functionals

Assume we have a function $V \ni f : \mathbb{R}^N \to \mathbb{R}$ that takes N-dimensional vectors and maps them to real numbers.

Functionals are a generalization of this concept. Formally we have $F : V \to \mathbb{R}$. We can write these using an integral as

$$F[f] = \int_\Omega g(f(\mathbf{x}), \ldots)\mathrm{d}\mathbf{x}.$$

The function g can depend not only on $f(x)$ but it's derivatives (or more generally some linear operators acting on f).

Example

In Pset 2 we had the energy

$$H[u, v] = \frac{1}{2} \int_0^1 v(t, x)^2 + u'(t, x)^2 \mathrm{d}x,$$

where $v = \partial_t u$.

Example

In Pset 2 we had the energy

$$H[u,v] = \frac{1}{2}\int_0^1 v(t,x)^2 + u'(t,x)^2 \mathrm{d}x,$$

where $v = \partial_t u$.

This would be an example of a functional that can be evaluated for any $u, v \in V$. In this case the function g depends on the time and the spatial derivatives of u.

Example

In Pset 2 we had the energy

$$H[u, v] = \frac{1}{2}\int_0^1 v(t, x)^2 + u'(t, x)^2 \mathrm{d}x,$$

where $v = \partial_t u$.

This would be an example of a functional that can be evaluated for any $u, v \in V$. In this case the function g depends on the time and the spatial derivatives of u.

Another example: consider a graph $\gamma = \{(x, f(x)) : x \in (0, 1)\}$. The length of the graph (curve) is given by

$$L_\gamma[f] = \int_0^1 \sqrt{1 + f'(x)^2} \mathrm{d}x.$$

Derivatives of functions

Consider a function $f : \mathbb{R}^N \to \mathbb{R}$. How can we make sense of the derivative ∇?

Derivatives of functions

Consider a function $f : \mathbb{R}^N \to \mathbb{R}$. How can we make sense of the derivative ∇?

We can define the derivative as a directional derivative. We have

$$f(\mathsf{x} + \epsilon \mathsf{y}) = f(\mathsf{x}) + \epsilon \mathsf{y} \cdot \nabla f(\mathsf{x}) + \mathcal{O}(\epsilon^2).$$

Derivatives of functions

Consider a function $f : \mathbb{R}^N \to \mathbb{R}$. How can we make sense of the derivative ∇?

We can define the derivative as a directional derivative. We have

$$f(\mathsf{x} + \epsilon \mathsf{y}) = f(\mathsf{x}) + \epsilon \mathsf{y} \cdot \nabla f(\mathsf{x}) + \mathcal{O}(\epsilon^2).$$

We can define the gradient of f at x by

$$\langle \mathsf{y}, \nabla f(\mathsf{x}) \rangle = \mathsf{y} \cdot \nabla f(\mathsf{x}) = \lim_{\epsilon \to 0} \frac{f(\mathsf{x} + \epsilon \mathsf{y}) - f(\mathsf{x})}{\epsilon}$$

for all $\mathsf{y} \in \mathbb{R}^N$.

Derivatives of functions

Consider a function $f : \mathbb{R}^N \to \mathbb{R}$. How can we make sense of the derivative ∇?

We can define the derivative as a directional derivative. We have

$$f(\mathbf{x} + \epsilon \mathbf{y}) = f(\mathbf{x}) + \epsilon \mathbf{y} \cdot \nabla f(\mathbf{x}) + \mathcal{O}(\epsilon^2).$$

We can define the gradient of f at $\mathbf{x}$ by

$$\langle \mathbf{y}, \nabla f(\mathbf{x}) \rangle = \mathbf{y} \cdot \nabla f(\mathbf{x}) = \lim_{\epsilon \to 0} \frac{f(\mathbf{x} + \epsilon \mathbf{y}) - f(\mathbf{x})}{\epsilon}$$

for all $\mathbf{y} \in \mathbb{R}^N$.

We can choose $\mathbf{y}$ to be some basis vectors $\hat{\mathbf{x}}_i$ for $i = 1, ..., N$ giving us the coordinate representation

$$\nabla f(\mathbf{x}) = (\partial_{x_1} f(\mathbf{x}), \partial_{x_2} f(\mathbf{x}), ..., \partial_{x_N} f(\mathbf{x})).$$

Derivatives of functionals

We can generalize the notion of the gradient of a function. Let's define

$$\left\langle \phi, \frac{\delta F[f]}{\delta f} \right\rangle = \lim_{\epsilon \to 0} \frac{F[f + \epsilon\phi] - F[f]}{\epsilon} = \left(\frac{\mathrm{d}F[f + \epsilon\phi]}{\mathrm{d}\epsilon} \right)_{\epsilon=0}$$

for all $\phi \in C_0^\infty$ (the space here is important).

Derivatives of functionals

We can generalize the notion of the gradient of a function. Let's define

$$\left\langle \phi, \frac{\delta F[f]}{\delta f} \right\rangle = \lim_{\epsilon \to 0} \frac{F[f + \epsilon\phi] - F[f]}{\epsilon} = \left(\frac{\mathrm{d}F[f + \epsilon\phi]}{\mathrm{d}\epsilon} \right)_{\epsilon=0}$$

for all $\phi \in C_0^\infty$ (the space here is important).

In a similar way to a directional derivative, this measures the change of the functional F at f in the direction of ϕ.

Derivatives of functionals

We can generalize the notion of the gradient of a function. Let's define

$$\left\langle \phi, \frac{\delta F[f]}{\delta f} \right\rangle = \lim_{\epsilon \to 0} \frac{F[f + \epsilon\phi] - F[f]}{\epsilon} = \left(\frac{\mathrm{d}F[f + \epsilon\phi]}{\mathrm{d}\epsilon} \right)_{\epsilon=0}$$

for all $\phi \in C_0^\infty$ (the space here is important).

In a similar way to a directional derivative, this measures the change of the functional F at f in the direction of ϕ.

Since the inner product is some integral

$$\langle f, g \rangle = \int_\Omega f(\mathrm{x}) g(\mathrm{x}) \mathrm{d}\mathrm{x},$$

we can immediately see that $\delta F[f]/\delta f$ is a function (or a distribution). It is called the variational derivative or functional derivative of F.

Example

Let's consider the same functional as before i.e.

$$H[u, v] = \frac{1}{2}\int_0^1 v(t, x)^2 + u'(t, x)^2 \mathrm{d}x.$$

Example

Let's consider the same functional as before i.e.

$$H[u,v] = \frac{1}{2}\int_0^1 v(t,x)^2 + u'(t,x)^2 \mathrm{d}x.$$

Let's calculate

$$H[u, v+\epsilon\phi] = \frac{1}{2}\int_0^1 (v+\epsilon\phi)^2 + (u')^2 \mathrm{d}x = \frac{1}{2}\int_0^1 v^2 + (u')^2 \mathrm{d}x + \frac{1}{2}\int_0^1 2\epsilon\phi v \mathrm{d}x + \mathcal{O}(\epsilon^2).$$

Example

Let's consider the same functional as before i.e.

$$H[u, v] = \frac{1}{2} \int_0^1 v(t, x)^2 + u'(t, x)^2 \mathrm{d}x.$$

Let's calculate

$$H[u, v + \epsilon\phi] = \frac{1}{2} \int_0^1 (v + \epsilon\phi)^2 + (u')^2 \mathrm{d}x = \frac{1}{2} \int_0^1 v^2 + (u')^2 \mathrm{d}x + \frac{1}{2} \int_0^1 2\epsilon\phi v \mathrm{d}x + \mathcal{O}(\epsilon^2).$$

We recognize the first term as $H[u, v]$ giving

$$H[u, v + \epsilon\phi] - H[u, v] = \epsilon \int_0^1 \phi v \mathrm{d}x + \mathcal{O}(\epsilon^2) = \epsilon \langle \phi, v \rangle + \mathcal{O}(\epsilon^2).$$

Example

Let's consider the same functional as before i.e.

$$H[u,v] = \frac{1}{2}\int_0^1 v(t,x)^2 + u'(t,x)^2 \mathrm{d}x.$$

Let's calculate

$$H[u,v+\epsilon\phi] = \frac{1}{2}\int_0^1 (v+\epsilon\phi)^2 + (u')^2 \mathrm{d}x = \frac{1}{2}\int_0^1 v^2 + (u')^2 \mathrm{d}x + \frac{1}{2}\int_0^1 2\epsilon\phi v \mathrm{d}x + \mathcal{O}(\epsilon^2).$$

We recognize the first term as $H[u,v]$ giving

$$H[u,v+\epsilon\phi] - H[u,v] = \epsilon\int_0^1 \phi v \mathrm{d}x + \mathcal{O}(\epsilon^2) = \epsilon\left\langle \phi, v\right\rangle + \mathcal{O}(\epsilon^2).$$

We can divide by ϵ and take the limit giving

$$\left\langle \phi, \frac{\delta H[u,v]}{\delta v}\right\rangle := \lim_{\epsilon\to 0}\frac{H[u,v+\epsilon\phi] - H[u,v]}{\epsilon} = \left\langle \phi, v\right\rangle.$$

$$\left\langle \phi, \frac{\delta H[u,v]}{\delta v} \right\rangle := \lim_{\epsilon \to 0} \frac{H[u, v + \epsilon\phi] - H[u,v]}{\epsilon} = \langle \phi, v \rangle \, .$$

$$\left\langle \phi, \frac{\delta H[u,v]}{\delta v} \right\rangle := \lim_{\epsilon \to 0} \frac{H[u, v+\epsilon\phi] - H[u,v]}{\epsilon} = \langle \phi, v \rangle \, .$$

Since this holds for all $\phi \in C_0^\infty$ we conclude that

$$\frac{\delta H[u,v]}{\delta v} = v$$

in a weak sense (almost everywhere). Note that in practice we collected all the terms proportional to ϵ. Also, this just gave us the usual derivative of $v^2/2$.

Another example

Let's define the potential energy part of the above energy as

$$U[u] = \frac{1}{2}\int_0^1 u'(x)^2 \mathrm{d}x.$$

Another example

Let's define the potential energy part of the above energy as

$$U[u] = \frac{1}{2}\int_0^1 u'(x)^2 \mathrm{d}x.$$

Let's calculate the functional derivative $\delta U/\delta u$. We have

$$U[u+\epsilon\phi] = \frac{1}{2}\int_0^1 \left(\partial_x(u+\epsilon\phi)\right)^2 \mathrm{d}x = \frac{1}{2}\int_0^1 (u')^2 \mathrm{d}x + \epsilon\int_0^1 u'\phi' \mathrm{d}x + \mathcal{O}(\epsilon^2).$$

Another example

Let's define the potential energy part of the above energy as

$$U[u] = \frac{1}{2}\int_0^1 u'(x)^2 \mathrm{d}x.$$

Let's calculate the functional derivative $\delta U/\delta u$. We have

$$U[u+\epsilon\phi] = \frac{1}{2}\int_0^1 \left(\partial_x(u+\epsilon\phi)\right)^2 \mathrm{d}x = \frac{1}{2}\int_0^1 (u')^2 \mathrm{d}x + \epsilon\int_0^1 u'\phi' \mathrm{d}x + \mathcal{O}(\epsilon^2).$$

Again, we recognize the first part as U and we have

$$\frac{U[u+\epsilon\phi] - U[u]}{\epsilon} = \int_0^1 u'\phi' \mathrm{d}x + \mathcal{O}(\epsilon).$$

We can use integration by parts for the first term on the right giving

$$\int_0^1 u'\phi' \mathrm{d}x = (\phi(x)u'(x))_{x=0}^1 - \int_0^1 \phi u'' \mathrm{d}x.$$

We can use integration by parts for the first term on the right giving

$$\int_0^1 u'\phi' \mathrm{d}x = (\phi(x)u'(x))_{x=0}^1 - \int_0^1 \phi u'' \mathrm{d}x.$$

The boundary term evaluates to 0 since $\phi \in C_0^\infty$. The remaining term can be written as an inner product giving

$$\lim_{\epsilon \to 0} \frac{U[u+\epsilon\phi] - U[u]}{\epsilon} =: \left\langle \phi, \frac{\delta U}{\delta u} \right\rangle = - \left\langle \phi, u'' \right\rangle .$$

We can use integration by parts for the first term on the right giving

$$\int_0^1 u'\phi' \mathrm{d}x = (\phi(x)u'(x))_{x=0}^1 - \int_0^1 \phi u'' \mathrm{d}x.$$

The boundary term evaluates to 0 since $\phi \in C_0^\infty$. The remaining term can be written as an inner product giving

$$\lim_{\epsilon \to 0} \frac{U[u+\epsilon\phi] - U[u]}{\epsilon} =: \left\langle \phi, \frac{\delta U}{\delta u} \right\rangle = - \left\langle \phi, u'' \right\rangle .$$

We conclude that

$$\frac{\delta U[u]}{\delta u} = -u''.$$

Notice that this is not the usual derivative of the integrand in the functional.

In general we have for a functional

$$F[f] = \int_{\Omega} g\left(\left(\prod_{i} \partial_{x_i}^{n_i}\right) f(\mathbf{x})\right) \mathrm{d}\mathbf{x}$$

In general we have for a functional

$$F[f] = \int_\Omega g\left(\left(\prod_i \partial_{x_i}^{n_i}\right) f(\mathsf{x})\right) \mathrm{d}\mathsf{x}$$

$$\frac{\delta F[f]}{\delta f}(\mathsf{x}) = (-1)^{\sum_i n_i}\left(\prod_i \partial_{x_i}^{n_i}\right)\left[\frac{\partial g}{\partial\left(\prod_i \partial_{x_i}^{n_i}\right) f(\mathsf{x})}\right]$$

and any linear combinations of these.

This formula is somewhat messy so let’s give a simpler example.

Example

Let

$$F[f] = \int_\Omega g\left(f(\mathsf{x}), \nabla f(\mathsf{x}), \Delta f(\mathsf{x})\right) \mathrm{d}\mathsf{x}.$$

Example

Let

$$F[f] = \int_\Omega g\left(f(\mathsf{x}), \nabla f(\mathsf{x}), \Delta f(\mathsf{x})\right) \mathrm{d}\mathsf{x}.$$

Now,

$$\frac{\delta F}{\delta f}(\mathsf{x}) = \frac{\partial g}{\partial f(\mathsf{x})} - \nabla \cdot \frac{\partial g}{\partial \nabla f(\mathsf{x})} + \Delta \frac{\partial g}{\Delta f(\mathsf{x})}.$$

Here

$$\frac{\partial g}{\partial \nabla f(\mathsf{x})} = \left(\frac{\partial g}{\partial(\partial_{x_1} f(\mathsf{x}))}, \frac{\partial g}{\partial(\partial_{x_2} f(\mathsf{x}))}, \dots, \frac{\partial g}{\partial(\partial_{x_N} f(\mathsf{x}))} \right)^T.$$

Using variational derivatives

Let's use the energy we defined before but in a higher dimension:

$$U[u] = \frac{1}{2} \int_{\Omega} \|\nabla u(\mathbf{x})\|^2 \, \mathrm{d}\mathbf{x}.$$

Using variational derivatives

Let's use the energy we defined before but in a higher dimension:

$$U[u] = \frac{1}{2} \int_{\Omega} \|\nabla u(\mathbf{x})\|^2 \, \mathrm{d}\mathbf{x}.$$

Calculating the variational derivative gives

$$\frac{\delta U}{\delta u}(\mathbf{x}) = -\nabla \cdot (\nabla u(\mathbf{x})) = -\Delta u(\mathbf{x}).$$

Notice that $\|\nabla u(\mathbf{x})\|^2 = \nabla u(\mathbf{x}) \cdot \nabla u(\mathbf{x})$.

Using variational derivatives

Let's use the energy we defined before but in a higher dimension:

$$U[u] = \frac{1}{2} \int_{\Omega} \|\nabla u(\mathsf{x})\|^2 \, \mathrm{d}\mathsf{x}.$$

Calculating the variational derivative gives

$$\frac{\delta U}{\delta u}(\mathsf{x}) = -\nabla \cdot (\nabla u(\mathsf{x})) = -\Delta u(\mathsf{x}).$$

Notice that $\|\nabla u(\mathsf{x})\|^2 = \nabla u(\mathsf{x}) \cdot \nabla u(\mathsf{x})$.

We can now define for example the heat equation as

$$\partial_t u(t, \mathsf{x}) = -\frac{\delta U}{\delta u}(t, \mathsf{x}) = \Delta u(t, \mathsf{x}).$$

Let's look at the change of energy in time $\partial_t U[u]$. It turns out we can use a sort of chain rule

$$\partial_t U[u] = \left\langle \dot{u}, \frac{\delta U}{\delta u} \right\rangle .$$

You can compare this chain rule to a chain rule $\partial_t g(\mathsf{x}(t)) = \dot{\mathsf{x}} \cdot \nabla g(\mathsf{x}(t))$.

Let's look at the change of energy in time $\partial_t U[u]$. It turns out we can use a sort of chain rule

$$\partial_t U[u] = \left\langle \dot{u}, \frac{\delta U}{\delta u} \right\rangle .$$

You can compare this chain rule to a chain rule $\partial_t g(\mathsf{x}(t)) = \dot{\mathsf{x}} \cdot \nabla g(\mathsf{x}(t))$.

Plugging in the time evolution gives

$$\partial_t U[u] = -\left\langle \frac{\delta U}{\delta u}, \frac{\delta U}{\delta u} \right\rangle = -\left\| \frac{\delta U}{\delta u} \right\|^2 \leq 0.$$

Let's look at the change of energy in time $\partial_t U[u]$. It turns out we can use a sort of chain rule

$$\partial_t U[u] = \left\langle \dot{u}, \frac{\delta U}{\delta u} \right\rangle .$$

You can compare this chain rule to a chain rule $\partial_t g(\mathsf{x}(t)) = \dot{\mathsf{x}} \cdot \nabla g(\mathsf{x}(t))$.

Plugging in the time evolution gives

$$\partial_t U[u] = -\left\langle \frac{\delta U}{\delta u}, \frac{\delta U}{\delta u} \right\rangle = -\left\| \frac{\delta U}{\delta u} \right\|^2 \leq 0.$$

We immediately see that the energy is non-increasing in time. Notice that for this calculation we only used the fact that

$$\partial_t u(t, \mathsf{x}) = -\frac{\delta U}{\delta u}$$

without specifying the form of U. It follows that this is true for *all* dynamics that can be written like this using some functional U!

For many physical problems we have similar results. E.g. for the wave equation we have

$$
\begin{aligned}
\dot{v} &= -\frac{\delta H}{\delta u} = -\frac{\delta U}{\delta u} = \Delta u \\
\dot{u} &= \frac{\delta H}{\delta v} = v.
\end{aligned}
$$

For many physical problems we have similar results. E.g. for the wave equation we have

$$\begin{aligned}\dot{v} &= -\frac{\delta H}{\delta u} = -\frac{\delta U}{\delta u} = \Delta u \\ \dot{u} &= \frac{\delta H}{\delta v} = v.\end{aligned}$$

It follows that we can replace $\dot{v}$ by $\ddot{u}$ in the first equation giving the wave equation. These sort of dynamics will conserve H in time (you can try this at home) and are called Hamiltonian (physics) or symplectic (mathematics).

For many physical problems we have similar results. E.g. for the wave equation we have

$$\dot{v} = -\frac{\delta H}{\delta u} = -\frac{\delta U}{\delta u} = \Delta u$$
$$\dot{u} = \frac{\delta H}{\delta v} = v.$$

It follows that we can replace $\dot{v}$ by $\ddot{u}$ in the first equation giving the wave equation. These sort of dynamics will conserve H in time (you can try this at home) and are called Hamiltonian (physics) or symplectic (mathematics).

The reason we introduced these techniques is that using such energies as measures for stability and other kind of sanity checks with numerics is extremely helpful, especially for non-linear systems.