Homework #6: Calculus of Variations

Solutions

March 8, 2019

1 Derive the Euler-Lagrange Equation from Directional Derivative

Given an energy functional $E(f)$ of the form

$$E(f) = \int_0^1 L\Big(f(x), \dot{f}(x), x\Big)\,dx$$

we showed in class that a necessary condition for a function $f(x)$ to be a minimizer of E is that it satisfies the Euler-Lagrange equation given by

$$L_f - \frac{d}{dx}L_{\dot{f}} = 0.$$

We derived this condition by showing that given any perturbation function $g(x)$, the derivative in ϵ of $E\big(f(x) + \epsilon g(x)\big)$ vanishes at ϵ=0 if this equation is satisfied. Rederive the same Euler-Lagrange equation by showing instead that it causes the following directional derivative

$$\lim_{\epsilon\to 0} \frac{E(f+\epsilon g) - E(f)}{\epsilon}$$

to vanish for any function ("direction") g. (Hint: use the Taylor series approximation of L.)

Solution:

Using the Taylor series expansion of L we obtain

$$\begin{aligned}
\lim_{\epsilon\to 0} \frac{E(f+\epsilon g) - E(f)}{\epsilon} &= \lim_{\epsilon\to 0} \int_0^1 \frac{L\left(f(x)+\epsilon g(x), \dot{f}(x)+\epsilon\dot{g}(x), x\right) - L\left(f(x), \dot{f}(x), x\right)}{\epsilon} dx \\
&= \lim_{\epsilon\to 0} \int_0^1 \frac{L\left(f(x), \dot{f}(x), x\right) + \epsilon g(x)\, L_f\left(f(x), \dot{f}(x), x\right) + \epsilon\dot{g}(x)\, L_{\dot{f}}\left(f(x), \dot{f}(x), x\right) + O(\epsilon^2) - L\left(f(x), \dot{f}(x), x\right)}{\epsilon} dx \\
&= \lim_{\epsilon\to 0} \int_0^1 \frac{\epsilon g(x)\, L_f\left(f(x), \dot{f}(x), x\right) + \epsilon\dot{g}(x)\, L_{\dot{f}}\left(f(x), \dot{f}(x), x\right) + O(\epsilon^2)}{\epsilon} dx \\
&= \lim_{\epsilon\to 0} \int_0^1 \left(g(x)\, L_f\left(f(x), \dot{f}(x), x\right) + \dot{g}(x)\, L_{\dot{f}}\left(f(x), \dot{f}(x), x\right) + O(\epsilon)\right) dx \\
&= \int_0^1 \left(g(x)\, L_f\left(f(x), \dot{f}(x), x\right) + \dot{g}(x)\, L_{\dot{f}}\left(f(x), \dot{f}(x), x\right)\right) dx
\end{aligned}$$

Now applying integration by parts, and assuming that the allowing perturbations $g(x)$ vanish at $x = 0$ and $x = 1$ (in other words, $g(0) = g(1) = 0$) we obtain the familiar right hand side

$$\int g\left(L_f - \frac{d}{dx}L_{\dot{f}}\right) dx$$

which will only vanish for all choices of g if the Euler Lagrange euation $L_f - \frac{d}{dx}L_{\dot{f}} = 0$ is satisfied.

2 Finite-dimensional versus infinite-dimensional gradient

Validate the interpretation of the left hand side of the Euler-Lagrange equation as an infinite dimensional gradient by showing that the components of the "standard" (finite dimensional) gradient of

$$\hat{E}(f_0, f_1, \ldots, f_N) = \sum_{n=0}^{N-1} L\left(f_n, \frac{f_{n+1} - f_n}{\triangle x}, x_n\right) \triangle x$$

converge (up to a common scalar factor of $1/N$) to the values of $L_f - \frac{\partial}{\partial x} L_{\dot{f}}$. Note that $\hat{E}$ is a finite approximation of $E(f)$ from problem 1 where the variables $f_0, f_1, \ldots, f_N$ represent samples of the unknown function f at the points $x_0, x_1, \ldots, x_n$. We will assume uniform spacing $\triangle x = 1/N$ between samples, so that $x_0 = 0,\ x_1 = \triangle x,\ x_2 = 2\triangle x,\ \ldots$. As we increase the number N of samples, we better and better approximate the function f, but the dimension of the resulting gradient vector for $\hat{E}$ becomes larger and larger. Thus, as $N \to \infty$, the gradient becomes infinite dimensional.

Solution: Taking the partial derivative of $\hat{E}$ with respect to one of its arguments f_i yields the finite dimensional gradient component as follows

$$\begin{aligned}
\hat{E}_{f_i} &= \Delta x \frac{\partial}{\partial f_i}\left(\cdots + L\left(f_{i-2}, \frac{f_{i-1} - f_{i-2}}{\Delta x}, x_{i-2}\right) + L\left(f_{i-1}, \frac{f_i - f_{i-1}}{\Delta x}, x_{i-i}\right) + L\left(f_i, \frac{f_{i+1} - f_i}{\Delta x}, x_i\right) + L\left(f_{i+1}, \frac{f_{i+2} - f_{i+1}}{\Delta x}, x_{i+1}\right) + \cdots\right) \\
&= \frac{1}{N}\left(\frac{\partial}{\partial f_i} L\left(f_{i-1}, \frac{f_i - f_{i-1}}{\Delta x}, x_{i-i}\right) + \frac{\partial}{\partial f_i} L\left(f_i, \frac{f_{i+1} - f_i}{\Delta x}, x_i\right)\right) \\
&= \frac{1}{N}\left(\frac{1}{\Delta x} L_{f'}\left(f_{i-1}, \frac{f_i - f_{i-1}}{\Delta x}, x_{i-i}\right) - \frac{1}{\Delta x} L_{f'}\left(f_i, \frac{f_{i+1} - f_i}{\Delta x}, x_i\right) + L_f\left(f_i, \frac{f_{i+1} - f_i}{\Delta x}, x_i\right)\right) \\
&= \frac{1}{N}\left(L_f\left(f_i, \frac{f_{i+1} - f_i}{\Delta x}, x_i\right) - \frac{L_{f'}\left(f_i, \frac{f_{i+1} - f_i}{\Delta x}, x_i\right) - L_{f'}\left(f_{i-1}, \frac{f_i - f_{i-1}}{\Delta x}, x_{i-i}\right)}{\Delta x}\right) \\
&= \frac{1}{N}\left(L_f\left(f(x_i), \frac{f(x_i + \Delta x) - f(x_i)}{\Delta x}, x_i\right) - \frac{L_{f'}\left(f(x_i), \frac{f(x_i+\Delta x) - f(x_i)}{\Delta x}, x_i\right) - L_{f'}\left(f(x_i - \Delta x), \frac{f(x_i) - f(x_i - \Delta x)}{\Delta x}, x_i - \Delta x\right)}{\Delta x}\right)
\end{aligned}$$

If we now ignore the common scalar factor of $\frac{1}{N}$ we obtain that the rescaled (but same direction) gradient of $\hat{E}$ has components

$$L_f\left(f(x_i), \frac{f(x_i + \Delta x) - f(x_i)}{\Delta x}, x_i\right) - \frac{L_{f'}\left(f(x_i), \frac{f(x_i+\Delta x) - f(x_i)}{\Delta x}, x_i\right) - L_{f'}\left(f(x_i - \Delta x), \frac{f(x_i) - f(x_i - \Delta x)}{\Delta x}, x_i - \Delta x\right)}{\Delta x}$$

If we let $\Delta x \to 0$ (which is the same as letting $N \to \infty$) we get

$$\begin{aligned}
&\lim_{\Delta x \to 0} \quad L_f\left(f(x_i), \frac{f(x_i + \Delta x) - f(x_i)}{\Delta x}, x_i\right) - \frac{L_{f'}\left(f(x_i), \frac{f(x_i+\Delta x) - f(x_i)}{\Delta x}, x_i\right) - L_{f'}\left(f(x_i - \Delta x), \frac{f(x_i) - f(x_i - \Delta x)}{\Delta x}, x_i - \Delta x\right)}{\Delta x} \\
&= \quad L_f\left(f(x_i), \lim_{\Delta x \to 0} \frac{f(x_i + \Delta x) - f(x_i)}{\Delta x}, x_i\right) - \frac{\partial}{\partial x} L_{f'}\left(f(x_i), \lim_{\Delta x \to 0} \frac{f(x_i + \Delta x) - f(x_i)}{\Delta x}, x_i\right) \\
&= \quad L_f\left(f(x_i), f'(x_i), x_i\right) - \frac{\partial}{\partial x} L_{f'}\left(f(x_i), f'(x_i), x_i\right)
\end{aligned}$$

3 Calculate the gradient descent diffusion equation

Given an initial noisy 2D image I at time $t = 0$, derive the image diffusion equation

$$I_t = ?$$

which decreases the following energy (which is very large for a noisy image) as fast as possible. In other words, write down its gradient descent flow.

$$E = \frac{1}{2} \iint e^{(I_x^2 + I_y^2)}\, dx\, dy$$

Solution:

Notice that this has the form of the Perona-Malik style energy $\int_\Omega c(\|\nabla I\|)\, d\Omega$ where $c(z) = \frac{1}{2}e^{z^2}$, and therefore $\dot{c}(z) = ze^{z^2}$. As such we may apply the general gradient descent Perona-Malik diffusion formula derived in class as follows.

$$\begin{aligned}
I_t &= \nabla \cdot \left(\frac{\dot{c}(\|\nabla I\|)}{\|\nabla I\|} \nabla I \right) = \nabla \cdot \left(\frac{\|\nabla I\| e^{\|\nabla I\|^2}}{\|\nabla I\|} \nabla I \right) = \nabla \cdot \left(e^{\|\nabla I\|^2} \nabla I \right) \\
&= e^{\|\nabla I\|^2} \nabla \cdot \nabla I + \nabla \left(e^{\|\nabla I\|^2} \right) \cdot \nabla I = e^{\|\nabla I\|^2} \Delta I + \left(e^{\|\nabla I\|^2} \nabla \|\nabla I\|^2 \right) \cdot \nabla I \\
&= e^{\|\nabla I\|^2} \left(\Delta I + \nabla (I_x^2 + I_y^2) \cdot \nabla I \right) \\
&= e^{\|\nabla I\|^2} \left(\Delta I + 2 \begin{bmatrix} I_x I_{xx} + I_y I_{xy} \\ I_y I_{xy} + I_y I_{yy} \end{bmatrix} \cdot \begin{bmatrix} I_x \\ I_y \end{bmatrix} \right) \\
I_t &= e^{\|\nabla I\|^2} \left(I_{xx}(1 + 2I_x^2) + 4 I_x I_y I_{xy} + I_{yy}(1 + 2I_y^2) \right)
\end{aligned}$$

Taking the direct approach instead, we note that the Lagrangian $L(I, I_x, I_y, x, y) = \frac{1}{2} e^{(I_x^2 + I_y^2)}$ depends only on I_x and I_y and as such, the general gradient descent formula taking right-hand side of the Euler-Lagrange equation can be written as follows.

$$\begin{aligned}
I_t &= \frac{\partial}{\partial x} L_{I_x} + \frac{\partial}{\partial y} L_{I_x} \\
&= \frac{\partial}{\partial x} \left(I_x e^{(I_x^2 + I_y^2)} \right) + \frac{\partial}{\partial y} \left(I_y e^{(I_x^2 + I_y^2)} \right) \\
&= \left(I_{xx} e^{(I_x^2 + I_y^2)} + 2 I_x (I_x I_{xx} + I_y I_{xy}) e^{(I_x^2 + I_y^2)} \right) + \left(I_{yy} e^{(I_x^2 + I_y^2)} + 2 I_y (I_x I_{xy} + I_y I_{yy}) e^{(I_x^2 + I_y^2)} \right) \\
&= e^{(I_x^2 + I_y^2)} \left((1 + 2I_x^2) I_{xx} + 4 I_x I_y I_{xy} + (1 + 2I_y^2) I_{yy} \right) \\
&= e^{\|\nabla I\|^2} \left((1 + 2I_x^2) I_{xx} + 4 I_x I_y I_{xy} + (1 + 2I_y^2) I_{yy} \right)
\end{aligned}$$