1 Energy Minimization in Image Segmentation

Many image analysis methods can be phrased as energy minimization problems. For example, to segment some object in an image $I$, we might look for a model $M$ that minimizes the energy:

$$E(M) = E_{\text{int}}(M) + E_{\text{ext}}(M, I).$$

The first term, $E_{\text{int}}(M)$, is the internal energy of the model. This energy represents the properties we want to enforce, such as, boundary smoothness, statistical shape priors (similar to what we did with ASMs), etc. The second term, $E_{\text{ext}}(M, I)$, is the external energy of the model. This energy determines the match of the model to the image $I$ (low energy represents a “good” match to the image). It may be based on edge strength along the boundary, regional intensities, textures, etc.

I’m being purposefully vague about what type of mathematical object the model $M$ is. It could be a curve, surface, level set, medial axis representation, etc. Similarly, the image $I$ might be 2D, 3D, grayscale, color, etc. However, it is often the case that our model $M$ is some sort of function or mapping (curves, surface, and level sets all fit this description). In this case the energy $E$ is a functional, i.e., a map that takes a function as input and returns a real number as output.

2 Euler-Lagrange Derivation

Let’s first consider 1D functions $f : [a, b] \to \mathbb{R}$, and energy functionals of the form

$$E(f) = \int_a^b L(t, f, f') \, dt,$$

where $f' = \frac{df}{dt}$. The expression $L(t, f, f')$ is called the Lagrangian. To minimize such an energy, we will take the “derivative” with respect to the function parameter $f$. The idea is that if $f$ is a minimum point of $E(f)$, then any small perturbation of $f$ should increase the energy. That is,

$$E(f + \epsilon h) > E(f),$$

where $h$ is an arbitrary function and $\epsilon$ is some small number. One important consideration is the boundary conditions of the perturbation $h$. Here we will assume that $h(a) = h(b) = 0$. This is a perturbation that keeps endpoints fixed. We can now take a derivative with respect to the $\epsilon$ parameter:
\[
\frac{d}{d\epsilon}E(f + \epsilon h)\bigg|_{\epsilon=0} = \frac{d}{d\epsilon} \int_a^b L(t, f + \epsilon h, (f + \epsilon h)') \, dt \bigg|_{\epsilon=0} \\
= \int_a^b \left( \frac{\partial L}{\partial f} h + \frac{\partial L}{\partial f'} h' \right) \, dt \quad \text{Chain Rule} \\
= \int_a^b \left( \frac{\partial L}{\partial f} h - \frac{d}{dt} \frac{\partial L}{\partial f'} h \right) \, dt + \left. \frac{\partial L}{\partial f'} h \right|_a^b \quad \text{Integration by Parts} \\
= \int_a^b \left( \frac{\partial L}{\partial f} - \frac{d}{dt} \frac{\partial L}{\partial f'} \right) h \, dt \quad \text{Using } h(a) = h(b) = 0
\]

Now for \( f \) to be a critical point of the energy \( E \), we will require that the above derivative be equal to zero. Because \( h \) could be any function with zero endpoints, it must be the case that everything else inside the integral, i.e., the term inside parentheses, must equal zero. This is according to the fundamental lemma of the calculus of variations. So, the condition for \( f \) being a critical point of the energy \( E \) is given by the Euler-Lagrange equation:

\[
\frac{\partial L}{\partial f} - \frac{d}{dt} \frac{\partial L}{\partial f'} = 0 \quad (2)
\]

### 3 Example: Minimizing Arclength

The above derivations of the Euler-Lagrange equation was for real-valued functions \( f : [a, b] \to \mathbb{R} \). However, notice that because of linearity, the same derivation will work for vector-valued functions, i.e., functions of the form, \( f : [a, b] \to \mathbb{R}^n \). The only difference is that the derivatives \( \frac{\partial L}{\partial f} \) and \( \frac{\partial L}{\partial f'} \) are \( \mathbb{R}^n \)-valued, and the multiplications in the chain rule step are replaced with inner products.

Consider a curve \( c : [a, b] \to \mathbb{R}^n \) and the arclength functional:

\[
E(c) = \int_a^b \|c'\| \, dt. \quad (3)
\]

To minimize arclength, we apply the Euler-Lagrange equation. Let \( s(t) \) denote the arclength parameter function for \( c \), i.e., \( \frac{ds}{dt} = \|c'(t)\| \).

\[
\frac{\partial L}{\partial f} - \frac{d}{dt} \frac{\partial L}{\partial f'} = -\frac{d}{dt} \frac{c'(t)}{\|c'(t)\|} \\
= -\frac{d}{ds} T(s(t)) \frac{ds(t)}{dt} \quad \text{using } T = \frac{c'}{\|c'\|} \\
= -\kappa(t) N(t) \|c'(t)\| \quad \text{using } \frac{d}{ds} T = \kappa N
\]

Consider a curve with fixed endpoints, \( c(a) = p \) and \( c(b) = q \). The above equation tells us that the curve with minimal arclength between these two points must have zero curvature. In other words, \( c \) must be a straight line. Now consider a closed curve, i.e., \( c(a) = c(b) \). The above equation gives us a gradient descent algorithm that will shrink the curve \( c \) as fast as possible by moving in the normal direction proportional to the curvature. This is known as the curvature flow.