PCMI Lecture #5

Examples 1: ROF, CE, H1, MS, and K functional generalizations

Now we look more carefully at some variational image analysis methods.

Rudin · Osher - Fatemi (1992)

$$F(u) \equiv \int_\Omega |\nabla u|\, dx + \lambda \int_\Omega |u-d|^2 dx$$

previously we saw that this is equivalent to the maximum liklihood method using Bayes rule, when $\underbrace{p(d|u) \cong e^{-\lambda\int|u-d|^2dx}}_{\text{likelihood function}}$, $\underbrace{p(u) \cong e^{-\int|\nabla u|dx}}_{\text{prior probability}}$

Now we consider this functional more carefully. Note first that care must be taken with the total variation term if we want to apply this to images with edges or jump discontinuities.

① $|\nabla u| = \sup\limits_{|v|\le 1} \nabla u \cdot v$ Legendre-Fenchel dual of $\vee$ is $\sqcup$

② for smooth u and smooth vectorfields v with support contained in Ω, we have

$$\int_\Omega \nabla u \cdot v\, dx = -\int_\Omega u(\nabla\cdot v)\, dx$$

③ if u is smooth, then we can find a smooth v, $|v|\le 1$

$\ni$ $\int \nabla u\cdot v\, dx$ is arbitrarily close to $\int_\Omega |\nabla u|\, dx$.

$$\Rightarrow \int_\Omega |\nabla u|\, dx = \sup_v \int_\Omega u(\nabla\cdot v)\, dx \qquad v \text{ in } C^1,\ |v|\le 1$$

④ <u>Define</u> $$\int_\Omega |\nabla u|\, dx \equiv \sup_{\substack{v\in C^1\\ |v|\le 1}} \int_\Omega u(\nabla\cdot v)\, dx$$

BV: u in $L^1_{loc}(\Omega)$ and for each $V \subset\subset \Omega$ $\int_V |\nabla u|\, dx < \infty$

We will delve into this in more detail in later lectures (including quite a bit more in the GMT lectures) but here is one more result on BV functions.

Structure Theorem for BV_{loc} Functions

Let $u \in BV_{loc}(\Omega)$. Then $\exists$ Radon μ and a μ-measurable function $\sigma : \Omega \to \mathbb{R}^n$ $\ni$

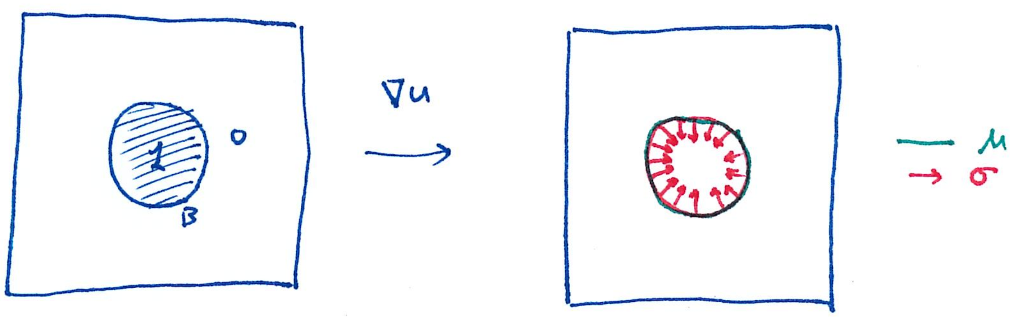
(i) $|\sigma(x)| = 1$ μ a.e. and

(ii) $\displaystyle\int_\Omega u \operatorname{div} v \, dx = -\int_\Omega v \cdot \sigma \, d\mu$

$\forall \; v \in C^1_c(\Omega; \mathbb{R}^n)$

Examples:

① $u = \chi_B$



$\mu = \mathcal{H}^1 \llcorner \partial B$

② u smooth

$$d\mu = |\nabla u| \, d\mathcal{L}^n \quad \text{i.e.} \quad \frac{d\mu}{d\mathcal{L}^n} = |\nabla u|$$

$$\sigma = \frac{\nabla u}{|\nabla u|} \quad \text{for } \{x \mid \nabla u \neq 0\}$$

③

Jump discontinuities at every dyadic vertical line: add a jump of size in the middle of each interval of length 2^{n-1} of size $\frac{1}{2^n \cdot 2^{n-1}}$.

The ROF functional has been extensively studied since its introduction in 1992. (I include all the things it has inspired)
In particular, it has inspired Total variation regularization of just about every image analysis task possible.

- inpainting
- deblurring
- super-resolution
- tomography

Algorithms for computing ROF minimizers have also proliferated.

① Gradient Descent

$$-\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda (u - d)$$

is the variational derivative of the ROF functional.

So, we simply go downhill!

$$u_t = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - \lambda (u-d)$$

This is regularized (to avoid division by zero) by ~~replacing~~ replacing $|\nabla u|$ with $\sqrt{\varepsilon + \nabla u \cdot \nabla u}$.

② Chambolle's Algorithm.

We need a few key facts and definitions from convex analysis

① $J^*(u) \equiv \sup_w \langle w, u \rangle - J(w)$

$J^{**}(v) \equiv \sup_w \langle w, v \rangle - J^*(w)$

② J convex, lsc $\rightarrow J^{**} = J$

expression for J^{**} is an envelope, in fact the envelope of support hyperplanes of J.

③ if J is one-homogeneous
$(J(\lambda u) = \lambda J(u) \quad \lambda > 0)$
then J^* is χ_K some closed convex set K.

(4) $w \in \partial J(u)$

$$\Updownarrow$$

$u \in \partial J^*(w)$

- argue from picture for one-homogeneous J's
- show analytic argument by observe that $w \in \partial J(u) \iff J^*(u) = \langle w, u\rangle - J(u)$

(5) write ROF as

$$\min_u \left\{ \frac{\|u - d\|^2}{2\lambda} + J(u) \right\}$$

$$\Downarrow$$

$$0 \in \{u - d + \lambda \partial J(u)\}$$

use (4) to get

$$u \in \partial J^*\left(\frac{d-u}{\lambda}\right)$$

$$\Downarrow$$

$$\frac{d}{\lambda} \in \frac{d-u}{\lambda} + \frac{1}{\lambda} \partial J^*\left(\frac{d-u}{\lambda}\right)$$

$$\Downarrow$$

$$0 \in \left(\frac{d-u}{\lambda} - \frac{d}{\lambda}\right) + \frac{1}{\lambda} \partial J^*\left(\frac{d-u}{\lambda}\right)$$

$\Rightarrow \frac{d-u}{\lambda}$ minimizes

$$*) \qquad \frac{\|w - d/\lambda\|}{2} + \frac{1}{\lambda} J^*(w)$$

But $J^*(w)$ is χ_K for a closed convex set... in fact that set are all the vectors that define supporting hyperplanes for $J(w)$.

looking Back at $\int |\nabla u|\,dx = \sup \int u(\nabla \cdot v)\,dx \quad |v| \le 1, \; v \in C_c^1(\Omega : \mathbb{R}^n)$

we see that K is $\{ \operatorname{div}(v) \mid |v| \le 1, \; v \in C_c^1(\Omega : \mathbb{R}^n)\}$

But minimizing $*)$ is simply finding the projection of d/λ onto K !!

(3) other algorithms: there are a fair number of others, but I will delay the presentation of the other one I will cover until later. That algorithm is the graph cut algorithm.

L^1TV, TVL^1, Chan-Esedoglu, ANCE

A seemingly small change:

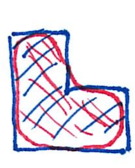
$$F(u) = \int |\nabla u| + \lambda \int |u-d|\, dx$$

produces significant changes in minimizers. We will study this L^1TV functional much more closely later, but ~~[illegible]~~ we will begin to explore it's properties here.

Euler-Lagrange: formally

$$-\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right) + \lambda \frac{u-d}{|u-d|} = 0$$

characterizes minimizers ... but this is singular for nonsmooth u and when $\nabla u = 0$. To handle this carefully we need to resort to geometric measure theory. But what this tells us is actually correct ... ~~[illegible]~~



consider the $u = \frac{1}{2}$ level set Σ

points on boundary of minimizers inside of Ω have negative curvature $= -\lambda$ on the outside ... $+\lambda$.

L^1TV minimizers

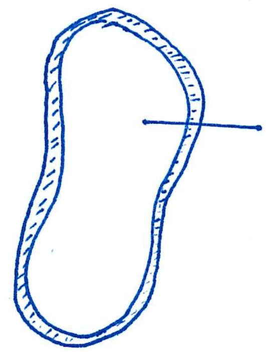
① $d = \chi_\Omega \Rightarrow \exists\, \Sigma \ni u = \chi_\Sigma$ is a minimizer

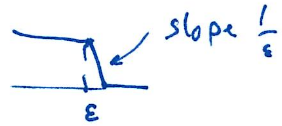
② $\int |\nabla u|\,dx + \lambda \int |u-d|\,dx$ is convex but <u>not</u> strictly convex.

$\Rightarrow$ minimal set is convex, but not necessarily a single point u.

③ for $u = \chi_\Sigma$ $\quad \int |\nabla u|\,dx = Per(\Sigma)$

for $u = \chi_\Sigma,\ d = \chi_\Omega$ $\quad \int |u-d|\,dx = |\Sigma \Delta \Omega|$

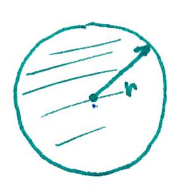




$\int |\nabla u|\,dx = \int |\nabla \chi_\varepsilon|\,dx = Per(\Sigma)$

④ Solutions of L^1TV can be generated by solving the set based problem for each level and stacking them up: we have among other things monotonicity.

Examples:

① 

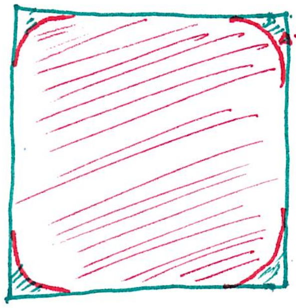
$\Omega = \chi_{B(0,r)}$

$\Sigma = B(0,r)$ for $\frac{2}{\lambda} < r$

$\Sigma = \{\alpha B(0,r)\}$ $\quad 0 \le \alpha \le 1$ $\quad \frac{2}{\lambda} = r$

$\Sigma = \phi$ $\quad \frac{2}{\lambda} < r$

②

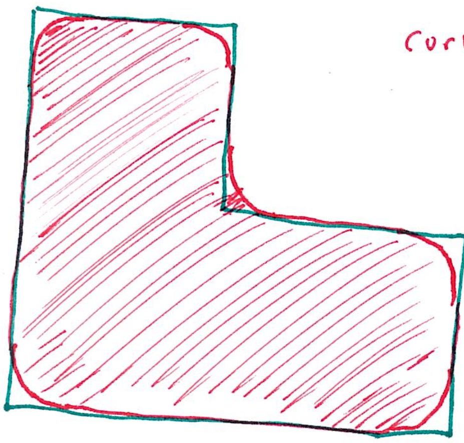


arc of radius $\frac{1}{\lambda}$

Σ minimizer

Ω input

③

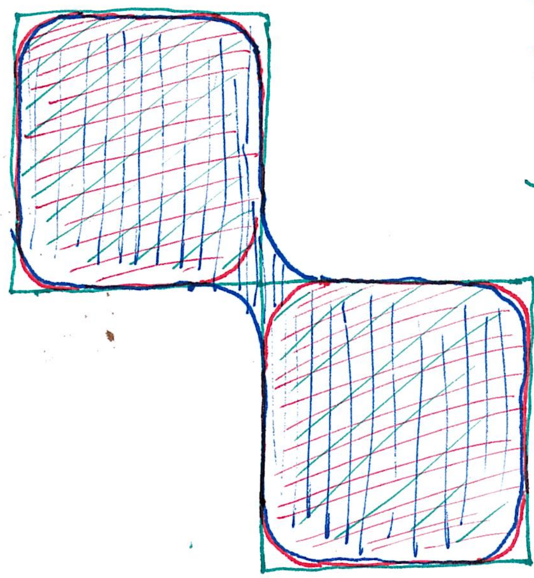


corners are arcs of radius $\frac{1}{\lambda}$

Σ minimizer

Ω input

④



$\left.\begin{array}{l}\Sigma \\ \Sigma\end{array}\right\}$ Both solutions minimizers

Ω input

Fact: $B_{2/\lambda} \subset \Omega \Rightarrow B_{2/\lambda} \subset \Sigma$

$B_{2/\lambda} \subset \Omega^c \Rightarrow B_{2/\lambda} \subset \Sigma^c$

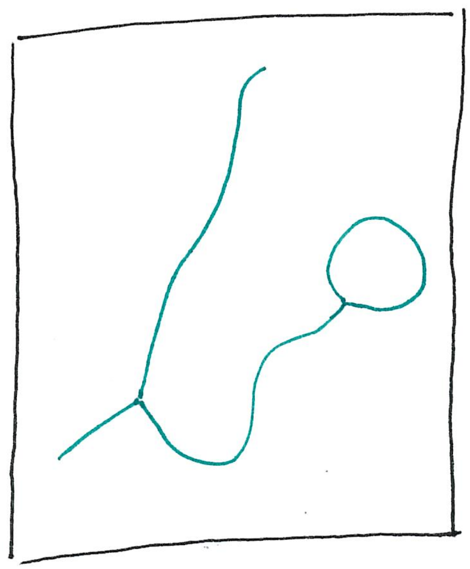
$$\underbrace{\Sigma_\alpha = \alpha \chi_{\Sigma} + (1-\alpha)\chi_{\Sigma}}_{\text{all solutions}}$$

Mumford-Shah Functional

$$F(u,\Gamma) \equiv \gamma_1 \int_{\Omega\setminus\Gamma} |\nabla u|^2 dx + \gamma_2 \mathcal{H}^{n-1}(\Gamma) + \gamma_3 \int_{\Omega} |u-d|^2 dx$$

$u : \Omega \subset \mathbb{R}^n \to \mathbb{R}$

Γ is a jump set.



$u \in SBV$ smooth ⊕ Jumps ... no Cantor like singularities

Mumford-Shah conjecture in the plane

Γ is a finite union of C^1 curves, that meet only at their ends, by sets of three and with only $120°$ angles.

Peetre's K-Functional and Scalespace from Heat Convolutions

$$\min_{u} \{ E_1(u) + t\, E_2(f-u) \}$$

choosing various E_1's / E_2's and a sequence of scales τ^n $0<\tau<1$ $n \in \mathbb{Z}$

we get a multiscale "expression" for f.

In particular, we can get something like this by simply running the heat equation with different stopping times, or convolving with the heat kernel with different widths.

$S_f(f)$ then yields important scales

$$S_t(f)(x) = \approx \frac{d}{dt}(h_t * f)(x)$$ (see Jones & Le)

Precisely: $$K_t(x) = \left(e^{-2\pi t |\xi|^2}\right)^{\vee}(x)$$

$$|Sf(x,t)| \equiv \left| t^{1-\alpha/2} \frac{\partial K_t}{\partial t} * f(x) \right|$$

use peaks of Sf to identify scales $\bar{t}$ then use

$$\inf_{u \in BV} K(u) \equiv \|u\|_{BV} + \lambda \| K_{\bar{t}} * (f-u) \|_{L^1}$$

to ~~[illegible]~~ control what pieces of the data we are matching