Lecture 13:

Image processing by minimization

(Variational approach)

1. Consider a minimization model (usually in continuous sense)

2. Derive a PDE related to minimization model

3. Discretize the PDE to get a linear system.

e.g. Total-variation (TV) denoising model. (ROF)

(Rudin - Osher - Fatemi)

Theorem:

$$\int_a^b \int_a^b \left(k(x,y)\, \nabla f(x,y) \right) \cdot \nabla g(x,y)\, dx\, dy = -\int_a^b \int_a^b \nabla \cdot \left(k(x,y)\, \nabla f(x,y) \right) g(x,y)\, dx\, dy$$

where $k : [a,b] \times [a,b] \to \mathbb{R}$.

Proof:

$$\int_a^b \int_a^b \left[k(x,y) \frac{\partial f}{\partial x} \frac{\partial g}{\partial x} + k(x,y) \frac{\partial f}{\partial y} \frac{\partial g}{\partial y} \right] dx\, dy$$

$$= -\int_a^b \int_a^b \frac{\partial}{\partial x}\left(k(x,y) \frac{\partial f}{\partial x} \right) g\, dx\, dy + \int_a^b \cancelto{0}{k(x,y) \frac{\partial f}{\partial x}\, g \Big|_{x=a}^{x=b}}\, dy$$

$$- \int_a^b \int_a^b \frac{\partial}{\partial y}\left(k(x,y) \frac{\partial f}{\partial y} \right) g\, dx\, dy + \int_a^b \cancelto{0}{k(x,y) \frac{\partial f}{\partial y}\, g \Big|_{y=a}^{y=b}}\, dx$$

$$= -\int_a^b \int_a^b \underbrace{\left[\frac{\partial}{\partial x}\left(k(x,y) \frac{\partial f}{\partial x} \right) + \frac{\partial}{\partial y}\left(k(x,y) \frac{\partial f}{\partial y} \right) \right]}_{\nabla \cdot (k(x,y)\, \nabla f)} g\, dx\, dy$$

In general, we have:

<u>Useful Tool:</u> (Integration by part)

$$\nabla \cdot \left(V_1(x,y), V_2(x,y) \right) = \frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y}$$

$$\int_\Omega \nabla f \cdot \nabla g \; dx\,dy = - \int_\Omega \left(\overbrace{\nabla \cdot}^{\text{divergence}} (\nabla f) \right) g \; dx\,dy + \int_{\partial \Omega} g \, (\nabla f \cdot \vec{n}) \, ds$$

where $\vec{n} = (n_1, n_2)$ = outward normal on the boundary.

or more generally,

$$\int_\Omega K(x,y) \, \nabla f(x,y) \cdot \nabla g(x,y) \, dx\,dy = - \int_\Omega \nabla \cdot \left(K(x,y) \nabla f(x,y) \right) g(x,y) \, dx\,dy + \int_{\partial \Omega} g(x,y) \left(K(x,y) \nabla f(x,y) \cdot \vec{n} \right) ds$$

Another useful fact:

If: $\int_\Omega T(x,y)\, v(x,y)\, dx\,dy = 0$ for all $v(x,y)$

then, we can conclude $T(x,y) = 0$ in Ω

Example: Suppose we have the following integral equation:

$$\int_a^b \int_a^b \big(f(x,y) - g(x,y)\big)\, v(x,y)\, dx\,dy + \int_a^b \int_a^b \nabla \cdot \nabla f(x,y)\, K(x,y)\, v(x,y)\, dx\,dy = 0$$

for all $v(x,y)$. Then: we have:

$$\int_a^b \int_a^b \Big[\big(f(x,y) - g(x,y)\big) + k(x,y)\, \nabla \cdot \nabla f(x,y)\Big] v(x,y)\, dx\,dy = 0$$

for all $v(x,y)$

We can conclude: $\big(f(x,y) - g(x,y)\big) + k(x,y)\, \nabla \cdot \nabla f(x,y) = 0$

for all $(x,y) \in [a,b] \times [a,b]$

Remark: More generally, if we do not enforce $f(x,y) = g(x,y) = 0$ on the boundary of the image domain Ω. Then, we do not enforce

$$v(x,y) = 0 \quad \text{on} \quad \partial\Omega.$$

In this case, we have:

$$0 = S'(0) = \int_\Omega 2\big(f(x,y) - g(x,y)\big)\, v(x,y)\, dx\,dy + \int_\Omega 2\, \nabla f(x,y) \cdot \nabla v(x,y)\, dx\,dy$$

$$= \int_\Omega 2\big(f(x,y) - g(x,y)\big)\, v(x,y)\, dx\,dy - \int_\Omega 2\, \Delta f(x,y)\, v(x,y)\, dx\,dy + \int_{\partial\Omega} \big(2\, \nabla f(x,y) \cdot \vec{n}(x,y)\big)\, v(x,y)\, ds$$

Overall, we get: $\int_\Omega (f - g - \Delta f)\, v\, dx\,dy - \int_{\partial\Omega} (\nabla f \cdot \vec{n})\, v\, ds = 0$ for all v

We conclude:

$$\begin{cases} f - g - \Delta f = 0 & \text{in } \Omega \\ \nabla f \cdot \vec{n} = 0 & \text{in } \partial\Omega \end{cases} \quad \text{(PDE)}$$

Example: Consider an image denoising model to find $f : \underbrace{[a,b]\times[a,b]}_{\Omega} \to \mathbb{R}$ that minimizes:

$$E(f) = \int_a^b\int_a^b (f(x,y)-g(x,y))^2 \, dxdy + \int_a^b\int_a^b |\nabla f(x,y)|^4 \, dxdy.$$

Suppose f minimizes $E(f)$. Assume $f(x,y) = g(x,y) = 0$ for all $(x,y) \in \partial D$. Find a partial differential equation that f must satisfy.

Solution: Suppose f minimizes $E(f)$. For any $v : D \to \mathbb{R}$ such that $v(x,y) = 0$ on ∂D, we have:

$$\begin{cases} f^\varepsilon \overset{\text{def}}{=} f + \varepsilon v \text{ is an image with} \\ f^\varepsilon(x,y) = f(x,y) + \varepsilon v(x,y) = 0 \text{ on } \partial D. \end{cases}$$

Consider $S : \mathbb{R} \to \mathbb{R}$ where $S(\varepsilon) \overset{\text{def}}{=} E(f^{\varepsilon}) = E(f + \varepsilon v)$

Then, $S(0) = E(f) =$ minimum of E. Thus, S attains minimum at $\varepsilon = 0$.

$\therefore \left.\dfrac{dS}{d\varepsilon}\right|_{\varepsilon=0} = 0$ for all $v : D \to \mathbb{R}$

Now,

$$0 = \left.\frac{dS}{d\varepsilon}\right|_{\varepsilon=0} = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \left(\int_D \big(f(x,y) + \varepsilon v(x,y) - g(x,y)\big)^2 \, dx\, dy + \int_D |\nabla(f + \varepsilon v)(x,y)|^4 \, dx\, dy \right)$$

$$\left(|\nabla f + \varepsilon \nabla v|^2\right)^2$$

$$= \big((\nabla f + \varepsilon \nabla v)\cdot(\nabla f + \varepsilon \nabla v)\big)^2$$

$$= \left(\nabla f \cdot \nabla f + 2\varepsilon \nabla f \cdot \nabla v + \varepsilon^2 \nabla v \cdot \nabla v\right)^2$$

$$= \left(|\nabla f|^2 + 2\varepsilon \nabla f \cdot \nabla v + \varepsilon^2 |\nabla v|^2\right)^2$$

$$\therefore \quad 0 = \frac{dS}{d\varepsilon}(0) = \int_D 2\left(f(x,y) + \varepsilon v(x,y) - g(x,y)\right) v(x,y)\Big|_{\varepsilon=0} \, dxdy$$

$$+ \int_D 2\left(|\nabla f|^2 + 2\varepsilon \nabla f \cdot \nabla v + \varepsilon^2 |\nabla v|^2\right)\left(2\nabla f \cdot \nabla v + 2\varepsilon |\nabla v|^2\right)\Big|_{\varepsilon=0} \, dxdy$$

$$\Rightarrow \quad 0 = \int_D 2\left(f(x,y) - g(x,y)\right) v(x,y) \, dxdy + \int_D 4\left(|\nabla f|^2\right) \nabla f \cdot \nabla v \, dxdy$$

$$= \int_D 2\left(f(x,y) - g(x,y)\right) v(x,y) \, dxdy - \int_D \left(4 \, \nabla \cdot \left(|\nabla f|^2 \nabla f\right)(x,y)\right) v(x,y) \, dxdy$$

$$+ \cancelto{0}{\int_{\partial D} \left(4\left(|\nabla f|^2\right) \nabla f(x,y) \cdot \vec{n}(x,y)\right) \underbrace{v(x,y)}_{=\,0} \, dxdy}$$

All together, we have:

$$0 = \int_D \left(2\left(f(x,y) - g(x,y)\right) - 4\, \nabla \cdot \left(|\nabla f(x,y)|^2 \, \nabla f(x,y) \right) \right) v(x,y)\, dx\, dy$$

for all $v(x,y)$.

We can conclude that:

$$f(x,y) - g(x,y) - 4\, \nabla \cdot \left(|\nabla f(x,y)|^2 \, \nabla f(x,y) \right) = 0 \quad \text{in } D.$$

(Partial differential equation)

Total variation (TV) denoising (ROF)

Invented by: Rudin, Osher, Fatemi

Motivation: Previous model: $f = g + \Delta f$. Solve for f from noisy g.

Disadvantage: smooth out edge.

Modification: $f = g + \nabla \cdot (K \nabla f)$

K is small on edges!!

Goal: Given a noisy image $g(x,y)$, we look for $f(x,y)$ that solves:

$$f = g + \lambda \frac{\partial}{\partial x}\left(\frac{1}{|\nabla f|(x,y)} \frac{\partial f}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{1}{|\nabla f|(x,y)} \frac{\partial f}{\partial y}\right) \qquad (*)$$

Remark: Problem arises if $|\nabla f(x,y)| = 0$. Take care of it later.

We'll show that $(*)$ must be satisfied by a minimizer of:

$$J(f) = \frac{1}{2}\int_\Omega (f(x,y) - g(x,y))^2 + \underbrace{\lambda}_{\text{constant parameter} > 0} \int_\Omega |\nabla f(x,y)|\, dx\, dy$$

Same idea: Let $S(\varepsilon) := E(f+\varepsilon v)$

$$= \frac{1}{2}\int_\Omega (f+\varepsilon v - g)^2 + \lambda \int_\Omega \underbrace{|\nabla f + \varepsilon \nabla v|}_{\sqrt{(\nabla f + \varepsilon \nabla v)\cdot(\nabla f + \varepsilon \nabla v)}}$$

$$\frac{d}{d\varepsilon} S(\varepsilon) = \left[\int_\Omega (f+\varepsilon v - g)\, v + \lambda \int_\Omega \frac{\nabla f \cdot \nabla v + 2\varepsilon\, \nabla v \cdot \nabla v}{\sqrt{(\nabla f + \varepsilon \nabla v)\cdot(\nabla f + \varepsilon \nabla v)}} \right]$$

If f is a minimizer, $\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} S(\varepsilon) = 0$ for all v.

$$\therefore\ S'(0) = 0 = \int_\Omega (f-g)\, v + \lambda \int_\Omega \frac{\nabla f \cdot \nabla v}{|\nabla f|}$$

$$= \int_\Omega (f-g)\, v - \lambda \int_\Omega \nabla \cdot \left(\frac{\nabla f}{|\nabla f|} \right) v + \lambda \int_{\partial\Omega} \left(\frac{\nabla f}{|\nabla f|} \cdot \vec{n} \right) v$$

$$= \int_\Omega \left[(f-g) - \lambda \nabla \cdot \left(\frac{\nabla f}{|\nabla f|} \right) \right] v + \lambda \int_{\partial\Omega} \left(\frac{\nabla f}{|\nabla f|} \cdot \vec{n} \right) v \quad \text{for all } v$$

We conclude: $(f-g) - \lambda \nabla \cdot \left(\frac{\nabla f}{|\nabla f|} \right) = 0$!!

In the discrete case,

$$J(f) = \frac{1}{2}\sum_{x=1}^{N}\sum_{y=1}^{N}(f(x,y)-g(x,y))^2+\lambda\sum_{x=1}^{N}\sum_{y=1}^{N}\sqrt{(f(x+1,y)-f(x,y))^2+(f(x,y+1)-f(x,y))^2}$$

J can be regarded as a multi-variable function depending on:
$f(1,1), f(1,2), \ldots, f(1,N), f(2,1), \ldots, f(2,N), \ldots, f(N,N)$.

If f is a minimizer, then $\frac{\partial J}{\partial f(x,y)} = 0$ for all (x,y).

$$\begin{aligned}\frac{\partial J}{\partial f(x,y)} &= (f(x,y)-g(x,y)) + \lambda\frac{2(f(x+1,y)-f(x,y))(-1)+2(f(x,y+1)-f(x,y))(-1)}{2\sqrt{(f(x+1,y)-f(x,y))^2+(f(x,y+1)-f(x,y))^2}}\\ &+\lambda\frac{2(f(x,y)-f(x-1,y))}{2\sqrt{(f(x,y)-f(x-1,y))^2+(f(x-1,y+1)-f(x-1,y))^2}}\\ &+\lambda\frac{2(f(x,y)-f(x,y-1))}{2\sqrt{(f(x+1,y-1)-f(x,y-1))^2+(f(x,y)-f(x,y-1))^2}} = 0\end{aligned}$$

By simplification:

$$f(x,y) - g(x,y) = \lambda \left\{ \frac{f(x+1,y) - f(x,y)}{\sqrt{(f(x+1,y) - f(x,y))^2 + (f(x,y+1) - f(x,y))^2}} - \frac{f(x,y) - f(x-1,y)}{\sqrt{(f(x,y) - f(x-1,y))^2 + (f(x-1,y+1) - f(x-1,y))^2}} \right\}$$

$$+ \lambda \left\{ \frac{f(x,y+1) - f(x,y)}{\sqrt{(f(x+1,y) - f(x,y))^2 + (f(x,y+1) - f(x,y))^2}} - \frac{f(x,y) - f(x,y-1)}{\sqrt{(f(x+1,y-1) - f(x,y-1))^2 + (f(x,y) - f(x,y-1))^2}} \right\}$$

$$= \frac{\frac{\partial f}{\partial x}\Big|_{(x,y)}}{|\nabla f|_{(x,y)}} \qquad = \frac{\frac{\partial f}{\partial x}\Big|_{(x-1,y)}}{|\nabla f|_{(x-1,y)}} \qquad = \frac{\frac{\partial f}{\partial y}\Big|_{(x,y)}}{|\nabla f|_{(x,y)}} \qquad = \frac{\frac{\partial f}{\partial y}\Big|_{(x,y-1)}}{|\nabla f|_{(x,y-1)}}$$

Discretization of $f - g = \lambda\, \nabla \cdot \left(\frac{\nabla f}{|\nabla f|} \right)$

How to minimise $J(f)$

We consider the problem of finding f that minimizes $J(f)$.

In the discrete case, J depends on $f(x,y)$ for $\begin{matrix} x=1,2,\dots,N \\ y=1,2,\dots,N \end{matrix}$.

Consider a time-dependent image $f(x,y;\underbrace{t}_{\text{time}})$. Assuming that $f(x,y;t)$ satisfies:

$$\frac{df(\cdot,\cdot;t)}{dt} = -\nabla J(f(\cdot,\cdot;t)) \qquad (**)$$

We can show that $J(f(\ ;t))$ decreases as t increases.

Note that:

$$\frac{d}{dt} J(f(\cdot,\cdot;t)) = \nabla J(f(\cdot,\cdot;t)) \cdot \frac{d}{dt} f(\cdot,\cdot;t) = -\nabla J(f(\cdot,\cdot;t)) \cdot \nabla J(f(\cdot,\cdot;t))$$

$$= -|\nabla J(f(\cdot,\cdot;t))|^2 \leq 0.$$

$\therefore$ $J(f(\cdot,\cdot;t))$ is decreasing as t increases!!

In the discrete case,

$$\frac{f^{n+1}-f^n}{\Delta t} = -\nabla J(f^n)$$

(Gradient descent algorithm)

For the ROF model:

$$\begin{aligned}
&\frac{f^{n+1}(x,y)-f^n(x,y)}{\Delta t} \\
&= -(f^n(x,y)-g(x,y)) + \lambda \frac{f^n(x+1,y)-f^n(x,y)}{\sqrt{(f^n(x+1,y)-f^n(x,y))^2+(f^n(x,y+1)-f^n(x,y))^2}} \\
&- \lambda \frac{f^n(x,y)-f^n(x-1,y)}{\sqrt{(f^n(x,y)-f^n(x-1,y))^2+(f^n(x-1,y+1)-f^n(x-1,y))^2}} \\
&+ \lambda \frac{f^n(x,y+1)-f^n(x,y)}{\sqrt{(f^n(x+1,y)-f^n(x,y))^2+(f^n(x,y+1)-f^n(x,y))^2}} \\
&- \lambda \frac{f^n(x,y)-f^n(x,y-1)}{\sqrt{(f^n(x+1,y-1)-f^n(x,y-1))^2+(f^n(x,y)-f^n(x,y-1))^2}}
\end{aligned}$$

Discretization of ∇J

(Gradient descent algorithm for ROF)

In the continuous case, consider:

$$E(f) = \int_\Omega (f-g)^2\,dx\,dy + \lambda \int |\nabla f|\,dx\,dy$$

We want to find a sequence $f_0 \overset{=}{} g, f_1, \ldots, f_n, \ldots$ such that:

$$E(f_0) \geq E(f_1) \geq \ldots \geq E(f_n) \geq E(f_{n+1}) \geq \ldots$$

Define $S(\varepsilon) \overset{def}{:=} E(f_n + \varepsilon v)$ for some suitable v. (Assuming that $\nabla f_n = 0$ on $\partial\Omega$)

For small $\varepsilon > 0$, by Taylor expansion,

$$\underbrace{S(\varepsilon)}_{E(f_{n+1})} = \underbrace{S(0)}_{E(f_n)} + S'(0)\underbrace{\varepsilon}_{>0} + \underbrace{\mathcal{O}(\varepsilon^2)}_{\text{neglible}}$$

If $S'(0) \leq 0$, then $E(f_{n+1}) \leq E(f_n)$

Now, $\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} S(\varepsilon) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \left[\int_\Omega (f_n + \varepsilon v - g)^2\,dx\,dy + \lambda \int_\Omega |\nabla f_n + \varepsilon \nabla v|\,dx\,dy \right]$

$$|\nabla f_n + \varepsilon \nabla v| = \sqrt{(\nabla f_n + \varepsilon \nabla v)\cdot(\nabla f_n + \varepsilon \nabla v)}$$

$$\therefore \quad S'(0) = \int_\Omega (f_n - g)\, v \, dx\,dy + \lambda \int_\Omega \frac{\nabla f_n \cdot \nabla v}{\underbrace{\sqrt{\nabla f_n \cdot \nabla f_n}}_{|\nabla f_n|}} \, dx\,dy$$

$$= \int_\Omega (f_n - g)\, v \, dx\,dy - \lambda \int_\Omega \nabla \cdot \left(\frac{\nabla f_n}{|\nabla f_n|} \right) v \, dx\,dy + \lambda \cancelto{0}{\int_{\partial\Omega} \frac{\nabla f_n}{|\nabla f_n|} \cdot \vec{n} \, v \, ds}$$

$$= \int_\Omega \left[(f_n - g) - \lambda \nabla \cdot \left(\frac{\nabla f_n}{|\nabla f_n|} \right) \right] v \, dx\,dy$$

Put $v = -\left[(f_n - g) - \lambda \nabla \cdot \left(\frac{\nabla f_n}{|\nabla f_n|} \right) \right]$. Then: $S'(0) \le 0$.

$\therefore$ Gradient descent algorithm:

$$f^{n+1} = f^n + \varepsilon \Bigg[\underbrace{- \left(2(f_n - g) - \lambda \nabla \cdot \left(\frac{\nabla f_n}{|\nabla f_n|} \right) \right)}_{v} \Bigg] \quad \text{for } n = 0, 1, 2, \ldots$$

This is called the gradient descent algorithm.