TV-norm minimization for images

① The continuum model of $u(x,y)$

$$\Omega = [0,1]\times[0,1]$$

$$L^2(\Omega) = \{ u(x,y) : \iint_\Omega |u(x,y)|^2\,dxdy < +\infty \}$$

$$H^1(\Omega) = \{ u(x,y) \in L^2(\Omega) : u_x, u_y \in L^2(\Omega) \}$$

$$\|u\|_{L^2} = \sqrt{\iint_\Omega |u|^2\,dxdy}$$

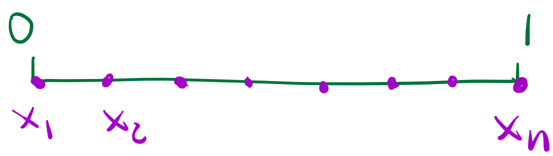
$$\nabla u = (u_x, u_y)$$

$$|\nabla u| = \sqrt{u_x^2 + u_y^2}$$

Define $\|u\|_{TV} = \iint_\Omega |\nabla u|\,dxdy = \iint_\Omega \sqrt{u_x^2+u_y^2}\,dxdy$

ROF model $\min\limits_{u\in H^1(\Omega)} \|u\|_{TV} + \lambda \|u-B\|_{L^2}^2$

② 1D Discrete Model



$\Delta x = \frac{1}{n-1}$

$u_j = u(x_j)$

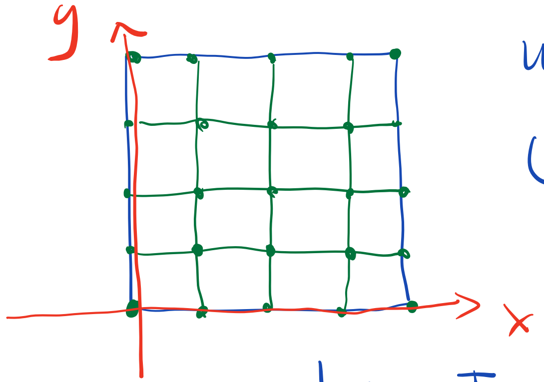
$$D = \begin{pmatrix} -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & -1 & 1 & & \\ & & & \ddots & \ddots & \\ & & & & -1 & 1 \\ & & & & & 0 \end{pmatrix}_{n\times n}. \qquad D^T = \begin{pmatrix} -1 & & & & & \\ 1 & -1 & & & & \\ & 1 & -1 & & & \\ & & \ddots & \ddots & & \\ & & & 1 & -1 & \\ & & & & 1 & 0 \end{pmatrix}_{n\times n}.$$

$$U = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \qquad \frac{1}{\Delta x} DU = \begin{pmatrix} \frac{u_2-u_1}{\Delta x} \\ \vdots \\ \frac{u_n-u_{n-1}}{\Delta x} \\ 0 \end{pmatrix} \approx \begin{pmatrix} u'(x_1) \\ \vdots \\ u'(x_{n-1}) \\ 0 \end{pmatrix}$$

$$D^T D = \begin{pmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{pmatrix}$$

$$\left(\frac{1}{\Delta x} D\right)^T \left(\frac{1}{\Delta x} D\right) U = \frac{1}{\Delta x^2} D^T D U \approx -u''(x)$$

③ 2D Discrete Model



uniform grid (x_i, y_j) $\Delta x = \Delta y = h$

U is a $n \times n$ 2D array with

$U(j,i) = u(x_i, y_j)$

$$\begin{cases} U_x = \frac{1}{h} U D^T \approx u_x \\ U_y = \frac{1}{h} D U \approx u_y \end{cases}$$

$U_x(j,i)^2 = |U(j,i+1) - U(j,i)|^2$

Discrete TV Norm

$$\|u\|_{TV} \approx \sum_i \sum_j h^2 \sqrt{u_x^2(x_i, y_j) + u_y^2(x_i, y_j)}$$

$$\Rightarrow \|U\|_{TV} = \sum_i \sum_j h^2 \sqrt{U_x^2(j,i) + U_y^2(j,i)}$$

$U \in \mathbb{R}^{n \times n}$

④ The linear operator K and its adjoint K^*

1D Continuum

$\Omega = [0,1]$

$L^2(\Omega) = \{u(x) : \int_0^1 u^2(x)\,dx < +\infty\}$

$H_0^1(\Omega) = \{u(x) : \int_0^1 u^2(x) + |u'(x)|^2 dx < +\infty,\ u(0) = u(1) = 0\}$

$$K : H_0^1(\Omega) \longrightarrow L^2(\Omega)$$
$$u(x) \longmapsto u'(x)$$

$$\langle u, v \rangle = \int_0^1 u(x)\, v(x)\, dx$$

$$\forall u,v \in H_0^1,\ \langle Ku, v \rangle = \int_0^1 u'(x)\, v(x)\, dx = -\int_0^1 u(x)\, v'(x)\, dx$$
$$= \langle u, K^* v \rangle$$

$$\Rightarrow K^* : H_0^1(\Omega) \to L^2(\Omega)$$
$$v(x) \longmapsto -v'(x)$$

1D Discrete

$$K : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$
$$u \longmapsto Du \approx u'(x)$$

$$D = \begin{pmatrix} -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & -1 & 1 & & \\ & & & \ddots & \ddots & \\ & & & & -1 & 1 \\ & & & & & 0 \end{pmatrix}$$

$$\langle Du, v \rangle = v^T D u$$
$$= (D^T v)^T u$$
$$= \langle u, D^T v \rangle$$

$$D^T = \begin{pmatrix} -1 & & & & & \\ 1 & -1 & & & & \\ & 1 & -1 & & & \\ & & \ddots & \ddots & & \\ & & & 1 & -1 & \\ & & & & 1 & 0 \end{pmatrix}$$

$$\Rightarrow K^* : \mathbb{R}^n \to \mathbb{R}^n$$
$$v \mapsto D^T v \approx -v'(x)$$

2D Continuum $\quad \Omega = [0,1] \times [0,1]$

$$L^2(\Omega) = \left\{ u(x,y) : \iint_\Omega |u(x,y)|^2\, dx\, dy < +\infty \right\}$$

$$H_0^1(\Omega) = \left\{ u(x,y) \in L^2(\Omega) : u_x, u_y \in L^2(\Omega),\ u|_{\partial\Omega} = 0 \right\}$$

$$K : H_0^1(\Omega) \longrightarrow L^2(\Omega) \otimes L^2(\Omega)$$
$$u \longmapsto (u_x, u_y) = \nabla u$$

$$\forall \vec{v} = (v_1, v_2) \in H^1(\Omega) \otimes H^1(\Omega)$$

$$\langle Ku, \vec{v}\rangle = \iint_\Omega Ku \cdot \vec{v}\, dxdy$$
$$= \iint_\Omega (u_x v_1 + u_y v_2)\, dxdy$$
$$= -\iint_\Omega u \frac{\partial}{\partial x} v_1 + u \frac{\partial}{\partial y} v_2 \, dxdy$$
$$= -\iint_\Omega u (\nabla \cdot \vec{v})\, dxdy$$
$$= \langle u, K^* \vec{v}\rangle$$

The adjoint of gradient is negative divergence

$$K^* : H^1 \otimes H^1 \longrightarrow L^2$$
$$\vec{v} \longmapsto -\nabla \cdot \vec{v} = -(v_1)_x - (v_2)_y$$

2D Discrete

$$U \in \mathbb{R}^{n\times n}$$
$$K : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}^{n\times n} \otimes \mathbb{R}^{n\times n}$$
$$U \longmapsto \left(\frac{1}{h} U D^T, \frac{1}{h} D U\right)$$

$$\langle KU, \vec{v}\rangle = \langle \frac{1}{h} U D^T, V_1\rangle + \langle \frac{1}{h} D U, V_2\rangle$$

$X, Y \in \mathbb{R}^{n\times n}$ $\quad \langle X, Y\rangle = \sum_i \sum_j X_{ij} Y_{ij} = tr(X^T Y)$
$= tr(Y^T X)$

$tr(ABC) = tr(CAB)$

$tr(AB) = tr(BA)$

$$= \frac{1}{h}\left[tr(V_1^T U D^T) + tr(V_2^T D U)\right]$$
$$= \frac{1}{h}\left[tr(U D^T V_1^T) + tr[(D^T V_2)^T U]\right]$$
$$= \frac{1}{h}\left[tr[(V_1 D)^T U] + tr[(D^T V_2)^T U]\right]$$
$$= \langle U, \frac{1}{h} V_1 D\rangle + \langle U, \frac{1}{h} D^T V_2\rangle$$
$$= \langle U, K^* \vec{v}\rangle$$

$$K^*: \mathbb{R}^{n\times n} \otimes \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}^{n\times n}$$

$$\vec{v} \longmapsto \frac{1}{h} v_1 D + \frac{1}{h} D^T v_2$$

$$\frac{1}{h} DU \approx u_y \qquad\qquad \approx -\nabla\cdot\vec{v}$$

$$\frac{1}{h} D^T U \approx -u_y$$

⑤ TV-denoising for 2D images

Given a noisy image $B \in \mathbb{R}^{n\times n}$, want to solve

$$\min_{U\in\mathbb{R}^{n\times n}} \|U\|_{TV} + \frac{\lambda}{2h}\|U-B\|_{L^2}^2$$

$$\|U-B\|_{L^2}^2 = \sum_i\sum_j h^2\cdot(U_{ij}-B_{ij})^2$$

$\lambda = 10\sim15$ is usually good for images

$$\Longleftrightarrow \min_U \sum_i\sum_j \left[\sqrt{(DU)_{ij}^2+(UD^T)_{ij}^2} + \frac{\lambda}{2}|U_{ij}-B_{ij}|^2\right] h$$

$$\Longleftrightarrow \min_U f(KU) + g(U) \qquad (P)$$

$$KU = (UD^T, DU)$$

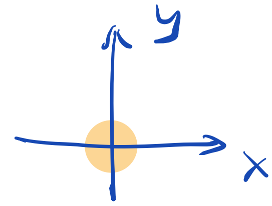
$$\vec{P} = (P_1, P_2)$$

$$g(U) = \frac{\lambda}{2}\sum_i\sum_j |U_{ij}-B_{ij}|^2$$

$f(\vec{P}) = \sum_i\sum_j \sqrt{(P_1)_{ij}^2+(P_2)_{ij}^2}$ is convex

$f(KU)$ is convex w.r.t. U

Example: $f(x,y) = \sqrt{x^2+y^2}$



So the correct subgradient is

$$\partial f(x,y) = \begin{cases} \nabla f = \left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}} \right) & \text{if } |x|+|y|>0 \\ \left(\left[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right], \left[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right] \right) & \text{if } x=y=0 \end{cases}$$

$$f^*(\vec{x}) = \max_{\vec{y}} \langle \vec{x}, \vec{y} \rangle - f(\vec{y}) \qquad \vec{y}_* = (y_1, y_2)$$

Critical point $\Rightarrow \vec{0} \in \vec{x} - \partial f(\vec{y}_*)$

$$\Rightarrow \vec{x} \in \partial f(\vec{y}_*)$$

$$\Rightarrow \begin{cases} \text{either } x_1 = \frac{y_1}{\sqrt{y_1^2+y_2^2}} \quad \text{or} \quad x_1 \in \left[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right] \\ \text{either } x_2 = \frac{y_2}{\sqrt{y_1^2+y_2^2}} \quad \text{or} \quad x_2 \in \left[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right] \end{cases}$$

$$\Rightarrow \text{either} \begin{pmatrix} x_1 = \frac{y_1}{\sqrt{y_1^2+y_2^2}} \\ x_2 = \frac{y_2}{\sqrt{y_1^2+y_2^2}} \end{pmatrix} \text{ or } \begin{pmatrix} x_1 \in \left[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right] \\ x_2 \in \left[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right] \end{pmatrix}$$

$$\Rightarrow f^*(\vec{x}) = \begin{cases} 0, & x_1^2+x_2^2 \le 1 \\ +\infty, & \text{otherwise} \end{cases}$$

So $\mathrm{Prox}^{\eta}_{f^*}$ is the projection to unit ball.

$$\mathrm{Prox}^{\eta}_{f^*}(x_1,x_2) = \begin{cases} \left(\dfrac{x_1}{\sqrt{x_1^2+x_2^2}}, \dfrac{x_2}{\sqrt{x_1^2+x_2^2}}\right), & x_1^2+x_2^2>1 \\ (x_1,x_2), & x_1^2+x_2^2\le 1 \end{cases}$$

2D TV Denoising

$$\min_{U} f(KU) + g(U) \qquad (P)$$

$$f(\vec{p}) = \sum_i \sum_j \sqrt{(p_1)_{ij}^2 + (p_2)_{ij}^2} \qquad g(U) = \frac{\lambda}{2}\sum_i\sum_j |U_{ij}-B_{ij}|^2$$

$$KU = (UD^T, DU) \approx \nabla u$$

$$K^*\vec{v} = -V_1 D - D^T V_2 \approx -\nabla\cdot\vec{v}$$

$$\min_{U\in\mathbb{R}^{n\times n}} f(KU) + g(U) \qquad (P)$$

$$\min_{U} \max_{\vec{v}\in[\mathbb{R}^{n\times n}]^2} \langle \vec{v}, KU\rangle - f^*(\vec{v}) + g(U) \qquad (PD)$$

$$-\min_{\vec{v}} f^*(\vec{v}) + g^*(-K^*\vec{v}) \qquad (D)$$

We have three algorithms

① (fast) prox-gradient on (D)

② (fast) PDHG

③ ADMM on (P)

PDHG is

$$\begin{cases} x_{k+1} = (I+\eta\partial g)^{-1}[x_k - \eta K^* y_k] \\ y_{k+1} = (I+\tau\partial f^*)^{-1}[y_k + \tau K(2x_{k+1}-x_k)] \end{cases}$$

$$\begin{cases} u_{k+1} = (I+\eta\partial g)^{-1}[u_k - \eta K^* \vec{v}_k] \\ \vec{v}_{k+1} = (I+\tau\partial f^*)^{-1}[\vec{v}_k + \tau K(2u_{k+1}-u_k)] \end{cases}$$

① If $\tau = \frac{1}{\eta}$ and $K = I$, it's Douglas-Rachford

② But $K \neq I$ here!

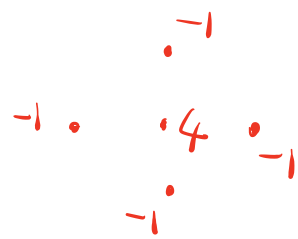
PDHG converges if $\eta\tau < \dfrac{1}{\underbrace{\rho(K^*K)}_{\text{largest eigenvalue magnitude of } K^*K}}$

$$\left.\begin{aligned} K &= \Delta \\ K^* &= -\nabla\cdot \end{aligned}\right\} \Rightarrow K^*K = -\Delta$$

$$KU = (UD^T, DU)$$

$$K^*U = -UD^TD - D^TDU \approx [-u_{xx} - u_{yy}]\, h^2$$

$$D^TD = \begin{pmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{pmatrix}$$



So K^*K is the 5-point discrete Laplacian and we know eigenvalues of $(K^*K) \geq 2\pi^2 \cdot h^2$ (c.f. Appendix in typed notes)

$$\Rightarrow \forall \eta > 0, \quad \tau < \frac{1}{2\pi^2 \eta \cdot h^2}$$

③ Difficult to implement Douglas-Rachford on (P) because no prox for $F(U) = f(KU)$

④ Douglas-Rachford on (D) $\Leftrightarrow$ ADMM on (P)

$$\begin{cases} z_{k+1} = \underset{z}{\operatorname{argmin}}\ g(z) + \frac{\eta}{2} \| Kz - \vec{w}_k + \frac{1}{\eta}\vec{v}_k \|^2 & \text{(a)} \\ \vec{w}_{k+1} = \underset{\vec{w}}{\operatorname{argmin}}\ f(\vec{w}) + \frac{\eta}{2} \| \vec{w} - \frac{1}{\eta}\vec{v}_k - Kz_{k+1} \|^2 & \text{(b)} \\ \vec{v}_{k+1} = \vec{v}_k + \eta (Kz_{k+1} - \vec{w}_{k+1}) & \text{(c)} \end{cases}$$

$$\text{(b)} \Leftrightarrow \vec{w}_{k+1} = \operatorname{Prox}_f^{1/\eta}\left[\frac{1}{\eta}\vec{v}_k + Kz_{k+1}\right]$$

Given $\vec{v} = (v_x, v_y) \in (\mathbb{R}^{n\times n}, \mathbb{R}^{n\times n})$

$$\operatorname{Prox}_f^{\frac{1}{\eta}}(\vec{v}) = (w_x, w_y), \quad P_{ij} = \sqrt{v_x^2(i,j) + v_y^2(i,j)}$$

$$\begin{cases} w_x(i,j) = \dfrac{v_x(i,j)}{p_{ij}}, \quad w_y(i,j) = \dfrac{v_y(i,j)}{p_{ij}} & \text{if } p_{ij} > 1 \\ (w_x, w_y) = (v_x, v_y) & \text{if } p_{ij} \leq 1 \end{cases}$$

For (a): Critical point

$$\Leftrightarrow \quad \nabla g(z_{k+1}) + \eta K^* [Kz - \vec{w}_k + \tfrac{1}{\eta}\vec{v}_k] = 0$$

$$\Leftrightarrow \quad \lambda[z_{k+1} - B] + \eta K^* K z_{k+1} + \eta K^*[-\vec{w}_k + \tfrac{1}{\eta}\vec{v}_k] = 0$$

$$\Leftrightarrow \quad \underbrace{[\lambda I + \eta K^* K]}\, z_{k+1} = \lambda B - \eta K^*[-\vec{w}_k + \tfrac{1}{\eta}\vec{v}_k]$$

$$[-\eta\Delta + \lambda I]\, z_{k+1} = \cdots$$

Efficient inversion of $[-\eta\Delta + \lambda I]$ can be easily done in Matlab/Python, code will be provided

---

Possible Final Project for undergrad & junior grad students:

Implement TV-minimization for image denoising

Sample codes will be provided.

Last Topic for Part II

For $\min_{x} \; f(x) + g(x)$

Two-operator Douglas-Rachford splitting (1979):

$$\begin{cases} y_{k+1} = \dfrac{I + R_f^{\eta} R_g^{\eta}}{2}(y_k) = y_k - x_k + \mathrm{Prox}_f^{\eta}(2x_k - y_k) \\ x_k = \mathrm{Prox}_g^{\eta}(y_k) \end{cases}$$

Convexity of f & g, $\forall \eta > 0$ $\Rightarrow$ Convergence

How about $\min_{x} \; f(x) + g(x) + h(x)$?

Three-operator splitting Davis-Yin (2016):

$$\begin{cases} y_{k+1} = \mathrm{Prox}_g^{\eta}(z_k) \\ x_{k+1} = \mathrm{Prox}_f^{\eta}(2y_{k+1} - z_k - \eta \nabla h(y_{k+1})) \\ z_{k+1} = z_k + x_{k+1} - y_{k+1} \end{cases}$$

$$\left.\begin{array}{l} \text{1) Convexity of } f, g, h \\ \text{2) } \nabla h \text{ is } L\text{-continuous} \\ \text{3) } \eta \le \frac{1}{L} \end{array}\right\} \Rightarrow \text{Convergence}$$