TV-norm minimization

① The continuum model

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

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

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

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

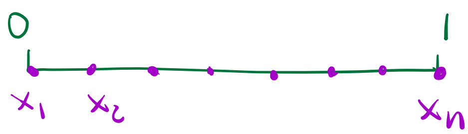
$$\nabla u = (u_x, u_y)$$

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

$$\|u\|_{H^1} = \sqrt{\iint_\Omega |u|^2 + |\nabla u|^2 \, dx\,dy} = \sqrt{\iint_\Omega u^2 + u_x^2 + u_y^2 \, dx\,dy}$$

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

② 1D Discrete Model



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

$$u_j = u(x_j)$$

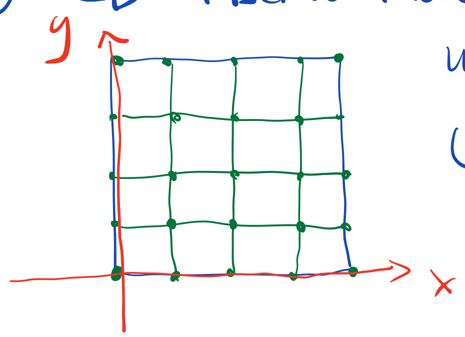
$$D = \begin{pmatrix} -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & -1 & 1 & & \\ & & & \ddots & \ddots & \\ & & & & -1 & 1 \\ & & & & & 0 \end{pmatrix}_{n \times n}. \qquad D^T = \begin{pmatrix} -1 & & & & & \\ 1 & -1 & & & & \\ & 1 & -1 & & & \\ & & \ddots & \ddots & & \\ & & & 1 & -1 & \\ & & & & 1 & 0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

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

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

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

③ 2D Discrete Model



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

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

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

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

Discrete TV Norm

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

$$\Rightarrow \|U\|_{TV} = \sum_i \sum_j h^2 \sqrt{\underbrace{U_x^2(j,i)}_{|U(j,i+1)-U(j,i)|^2} + U_y^2(j,i)}$$

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

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

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

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

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

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

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

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

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

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

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

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

$$\langle Ku, \vec{v}\rangle = \iint_\Omega Ku \cdot \vec{v}\, dx\,dy$$

$$= \iint_\Omega (u_x v_1 + u_y v_2)\, dx\,dy$$

$$= -\iint_\Omega u \tfrac{\partial}{\partial x} v_1 + u \tfrac{\partial}{\partial y} v_2 \, dx\,dy$$

$$= -\iint_\Omega u (\nabla \cdot \vec{v})\, dx\,dy$$

$$= \langle u, K^* \vec{v}\rangle$$

$$K^* : H^1 \otimes H^1 \longrightarrow L^2$$

$$\vec{v} \longmapsto -\nabla \cdot \vec{v} = -(v_1)_x - (v_2)_y$$

2D Discrete

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

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

$$U \longmapsto \left(\tfrac{1}{h} U D^T, \tfrac{1}{h} D U\right)$$

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

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

$$= tr(Y^T X)$$

$tr(ABC) = tr(CAB)$

$tr(AB) = tr(BA)$

$$= \tfrac{1}{h}\left[tr(V_1^T U D^T) + tr(V_2^T D U)\right]$$

$$= \tfrac{1}{h}\left[tr(U D^T V_1^T) + tr\left[(D^T V_2)^T U\right]\right]$$

$$= \tfrac{1}{h}\left[tr\left[(V_1 D)^T U\right] + tr\left[(D^T V_2)^T U\right]\right]$$

$$= \langle U, \tfrac{1}{h} V_1 D\rangle + \langle U, \tfrac{1}{h} D^T V_2\rangle$$

$$= \langle U, K^* \vec{V}\rangle$$

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

$$\vec{v} \longmapsto \frac{1}{h} v_1 D + \frac{1}{h} D^T v_2 \qquad \approx -\nabla \cdot \vec{v}$$

$$\frac{1}{h} D U \approx u_y$$

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

⑤ TV-denoising for 2D images

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

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

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

$\lambda = 10 \sim 15$ is usually good for images

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

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

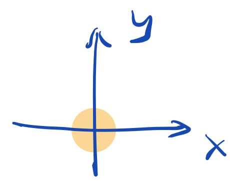
$$KU = (UD^T, DU)$$

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

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

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

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



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

$$\partial_x f = \begin{cases} \dfrac{x}{\sqrt{x^2+y^2}}, & \text{if } |x| + |y| > 0 \\ [-1, 1], & \text{if } x = y = 0 \ ? \end{cases}$$

$$\partial_y f = \begin{cases} \frac{y}{\sqrt{x^2+y^2}}, & \text{if } |x|+|y|>0 \\ [-1,1], & \text{if } x=y=0 \ ? \end{cases}$$

Recall subgradient / subdifferential is defined as

$$f(\vec{y}) \geq f(\vec{x}) + \langle \vec{z}, \vec{y}-\vec{x} \rangle, \quad \forall \vec{z} \in \partial f(\vec{x})$$

At $\vec{x} = (0,0)$, $\vec{f}(\vec{y}) \geq \langle \vec{z}, \vec{y} \rangle$

$$\iff \sqrt{y_1^2+y_2^2} \geq z_1 y_1 + z_2 y_2, \quad \forall \vec{y} \in \mathbb{R}^2$$

$$\iff (1-z_1^2) y_1^2 + (1-z_2^2) y_2^2 - 2 z_1 z_2 y_1 y_2 \geq 0$$

$$\iff (z_1 y_1 - z_2 y_2)^2 + (1-2z_1^2) y_1^2 + (1-2z_2^2) y_2^2 \geq 0$$

Obviously $z_1 = z_2 = 1$ does not work!

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

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

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

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

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

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

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

2D TV Denoising

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

$$f(\vec{P}) = \sum_i \sum_j \sqrt{(P_1)_{ij}^2 + (P_2)_{ij}^2} \qquad g(U) = \frac{\lambda}{2} \sum_i \sum_j |U_{ij} - B_{ij}|^2$$

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

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

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

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

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

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

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

$KU = (UD^T, DU)$

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

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

$$\begin{array}{ccc} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{array}$$

So K^*K is the 5-point discrete Laplacian and we know $\rho(K^*K)$ is at most 8

(c.f. my MA615 notes)

$\Rightarrow \forall \eta > 0, \ \tau < \frac{1}{8\eta}$

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)

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

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

For (a): Critical point

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

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

$$\Longleftrightarrow \underbrace{\left[\lambda I + \eta K^* K\right]}_{} z_{k+1} = \lambda B - \eta K^* \left[-\vec{w}_k + \frac{1}{\eta}\vec{v}_k\right]$$

$$\left[-\eta \Delta + \lambda I\right] z_{k+1} = \ldots$$