- Statistic interpretation of the optimization objective function

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}\in\mathbb{R}^n}{\arg\min}\left\{\frac{1}{2}\|\boldsymbol{y}-\boldsymbol{x}\|_2^2 + \lambda \mathrm{TV}(\boldsymbol{x})\right\}$$

True image

$$\boxed{\begin{array}{l} y = x+e. \\ e \sim N(0, \sigma^2). \end{array}}$$

observation : y. $\Rightarrow$ goal: get x from y.

true : x

method $x = \underset{x}{\operatorname{argmin}} \{f(x)\}$, $f(x) = \frac{1}{2}\|x-y\|^2 + \lambda r(x)$.

we show below why we use this objective function from a statistic perspective:

MAP(maximum a posteriori probability):

$$\hat{x}_{MAP} = \underset{x}{\operatorname{argmax}}\ P(x|y) = \underset{x}{\operatorname{argmax}} \frac{P(y|x)P(x)}{P(y)} = \underset{x}{\operatorname{argmax}}\ P(y|x)P(x).$$

$$= \underset{x}{\operatorname{argmin}}\ -\log\{P(y|x)\} - \log\{P(x)\}.$$

① $-\log\{P(y|x)\}$: log-likelihood.

② $-\log\{P(x)\}$: prior

PDF of Normal distribution: $\frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{(x-\mu)^2}{2\sigma^2})$

In our denoising model: $y = e+x$, $e \sim N(0, \sigma^2)$.

$\Rightarrow$ $y \sim N(x, \sigma^2)$. $\Rightarrow P(y|x) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{(y-x)^2}{2\sigma^2})$.

$$\begin{cases} ① -\log\{P(y|x)\} = C + \frac{1}{2\sigma^2}(y-x)^2. \\ ② \lambda r(x) \triangleq -\log P(x)\} \end{cases}$$

$$\hat{x}_{MAP} = \underset{x}{\operatorname{argmin}} \frac{1}{2\sigma^2}\|y-x\|^2 + \lambda r(x). = \underset{x}{\operatorname{argmin}} \underbrace{\frac{1}{2}\|y-x\|^2 + \lambda' r(x)}_{\text{Our objective function.}} = \underset{x}{\operatorname{argmin}}\ f(x).$$

- Optimization objective function

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}\in\mathbb{R}^n} \left\{ \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{x}\|_2^2 + \lambda \boxed{\mathrm{TV}(\boldsymbol{x})} \right\}$$

1: Anisotropic (our topic today)

$$\mathrm{TV}_1(\boldsymbol{x}) = \|D(\boldsymbol{x})\|_1 = \||D_x(\boldsymbol{x})| + |D_y(\boldsymbol{x})|\|_1$$

2: Isotropic (HW4 is based on this one)

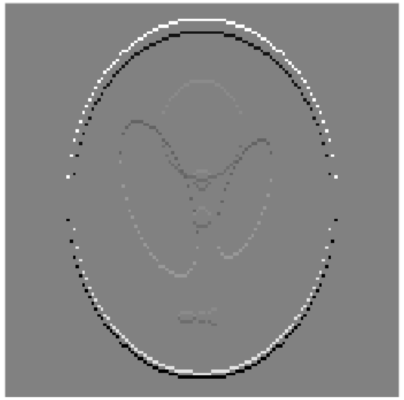
$$\mathrm{TV}_2(\boldsymbol{x}) = \|D(\boldsymbol{x})\|_2 = \|\sqrt{|D_x(\boldsymbol{x})|^2 + |D_y(\boldsymbol{x})|^2}\|_1$$

$\boldsymbol{x}$

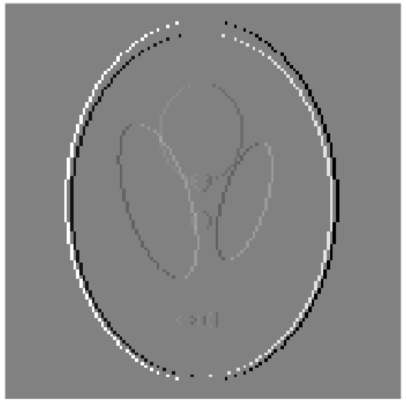
$D(\boldsymbol{x})$

$D_x(\boldsymbol{x})$ $D_y(\boldsymbol{x})$

Horizontal gradient



Vertical gradient



1: Anisotropic (our topic today)

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}\in\mathbb{R}^n} \left\{ \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{x}\|_2^2 + \lambda \mathrm{TV}(\boldsymbol{x}) \right\} \quad \text{with} \quad \mathrm{TV}_1(\boldsymbol{x}) = \|D(\boldsymbol{x})\|_1 = \| |D_x(\boldsymbol{x})| + |D_y(\boldsymbol{x})| \|_1$$

hints

Solution for TV-reg:

$P: \min\{\frac{1}{2}\|x-y\|_2^2 + \tau\|Dx\|_1\}$.

(ani) $TV_1 = \|Dx\|_1 = \|D_x x\|_1 + \|D_y x\|_1 = \sum_n (|[D_x x]_n| + |[D_y x]_n|) \Rightarrow \sum |\boxed{D_x x}| + |\boxed{D_y x}|$

(iso) $TV_2 = \|[Dx]_n\|_2 = \sum_n \sqrt{|[D_x x]_n|^2 + |[D_y x]_n|^2} \Longrightarrow \sum_n \sqrt{|\boxed{D_x x}|^2 + |\boxed{D_y x}|^2}$

↳各向同性

Property 1:

$$\frac{1}{p} + \frac{1}{q} = 1$$

$$\|x\|_p = \sup_{\|d\|_q \le 1} \{x^T d\} \qquad \text{(dual problem).}$$

i.e. ① $\|x\|_2 = \sup_{\|d\|_2 \le 1} \{x^T d\} \Rightarrow \begin{cases} d^* = \frac{x}{\|x\|_2} \\ \|x\|_2 = \sup\{x^T d^*\} = \|x\| \end{cases}$

② $\|x\|_1 = \sup_{\|d\|_\infty \le 1} \{x^T d\}. \Rightarrow \begin{cases} d^* = \mathrm{sign}\{x^T\}. \\ \|x\|_1 = \sum |x_i| \end{cases}$

We use this property to transfer ous problem to its dual problem ⇒.

1: Anisotropic (our topic today)

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}\in\mathbb{R}^n}{\arg\min} \left\{ \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{x}\|_2^2 + \lambda \mathrm{TV}(\boldsymbol{x}) \right\} \quad \text{with} \quad \mathrm{TV}_1(\boldsymbol{x}) = \|D(\boldsymbol{x})\|_1 = \||D_x(\boldsymbol{x})| + |D_y(\boldsymbol{x})|\|_1$$

hints

$$P: \min_x \left\{ \tfrac{1}{2}\|x-y\|^2 + \tau\|Dx\|_1 \right\} = \min_x \left\{ \tfrac{1}{2}\|x-y\|^2 + \tau \max_{\|d\|_\infty\le 1} \{[Dx]^T d\} \right\}$$

$$= \min_x \max_{\|d\|_\infty\le 1} \left\{ \tfrac{1}{2}\|x-y\|^2 + \tau[Dx]^T d \right\} = \max_{\|d\|_\infty\le 1} \min_x \left\{ \tfrac{1}{2}\|x-y\|^2 + \tau x^T D^T d \right\}$$

Define $f(x,d) \triangleq \frac{1}{2}\|x-y\|^2 + \tau x^T D^T d$.

$$\begin{aligned} f(x,d) &= \tfrac{1}{2}\|x-y\|^2 + \tau x^T D^T d \\ &= \tfrac{1}{2}\|x\|^2 - x^T y + \tfrac{1}{2}\|y\|^2 + x^T(\tau D^T d) \\ &= \tfrac{1}{2}\|x\|^2 - x^T(y-\tau D^T d) + \tfrac{1}{2}\|y-\tau D^T d\|^2 - \tfrac{1}{2}\|y-\tau D^T d\|^2 + \tfrac{1}{2}\|y\|^2 \\ &= \tfrac{1}{2}\|x-(y-\tau D^T d)\|^2 - \tfrac{1}{2}\|y-\tau D^T d\|^2 + \tfrac{1}{2}\|y\|^2 \end{aligned}$$

$$P: \min_x \left\{ \tfrac{1}{2}\|x-y\|^2 + \tau\|Dx\|_1 \right\} = \max_{\|d\|_\infty\le 1} \min_x \{ f(x,d) \}$$

$$= \max_{\|d\|_\infty\le 1} \min_x \left\{ \tfrac{1}{2}\|x-(y-\tau D^T d)\|^2 - \tfrac{1}{2}\|y-\tau D^T d\|^2 + \tfrac{1}{2}\|y\|^2 \right\}$$

$$= \max_{\|d\|_\infty\le 1} \min_x \left\{ \underbrace{\tfrac{1}{2}\|x-(y-\tau D^T d)\|^2}_{\text{quadratic}} - \underbrace{\tfrac{1}{2}\|y-\tau D^T d\|^2}_{\text{has nothin to do with } x.} \right\}.$$

We solve these two optimization problem separately $\Rightarrow$

1: Anisotropic (our topic today)

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}\in\mathbb{R}^n}{\arg\min} \left\{ \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{x}\|_2^2 + \lambda \mathrm{TV}(\boldsymbol{x}) \right\} \quad \text{with} \quad \mathrm{TV}_1(\boldsymbol{x}) = \|D(\boldsymbol{x})\|_1 = \| |D_x(\boldsymbol{x})| + |D_y(\boldsymbol{x})| \|_1$$

hints

For "min" part: $x^* = y - \tau D^T d^*$.

For "max" part: $d^* = \underset{\|d\|_\infty \le 1}{\arg\max} \left\{ -\frac{1}{2}\|y - \tau D^T d\|^2 \right\} = \underset{\|d\|_\infty \le 1}{\arg\min} \left\{ \frac{1}{2}\|y - \tau D^T d\|^2 \right\}$

This is a projection problem, and can be solved iteratively. using proximal gradient method.

$$d^+ = \mathrm{proj}_{\|\cdot\|_\infty \le 1}\left(d - r\tau D(\tau D^T d - y)\right), \text{ where } \mathrm{proj}_{\|\cdot\|_\infty}(z) = \begin{cases} z_i & \text{if } |z_i| \le 1 \\ \frac{z_i}{|z_i|} & \text{if } |z_i| > 1 \end{cases}$$

step size in the gradient descent step

To conclude:

problem:

$$P: \min_x \left\{ \frac{1}{2}\|x-y\|^2 + \tau\|Dx\|_1 \right\}$$

$$= \max_{\|d\|_\infty} \min_x \left\{ \frac{1}{2}\|x - (y - \tau D^T d)\|^2 - \frac{1}{2}\|y - \tau D^T d\|^2 \right\}$$

Solution:

For $t = 1, 2, \ldots$

$$d^t = \mathrm{proj}_{\|\cdot\|_\infty \le 1}\left(d^{t-1} - r\tau D(\tau D^T d^{t-1} - y)\right)$$

$$x^t = y - \tau D^T d^t$$

This is your reconstructed image at iteration t.