Hessian-Based Norm Regularization for Image Restoration with Biomedical Applications

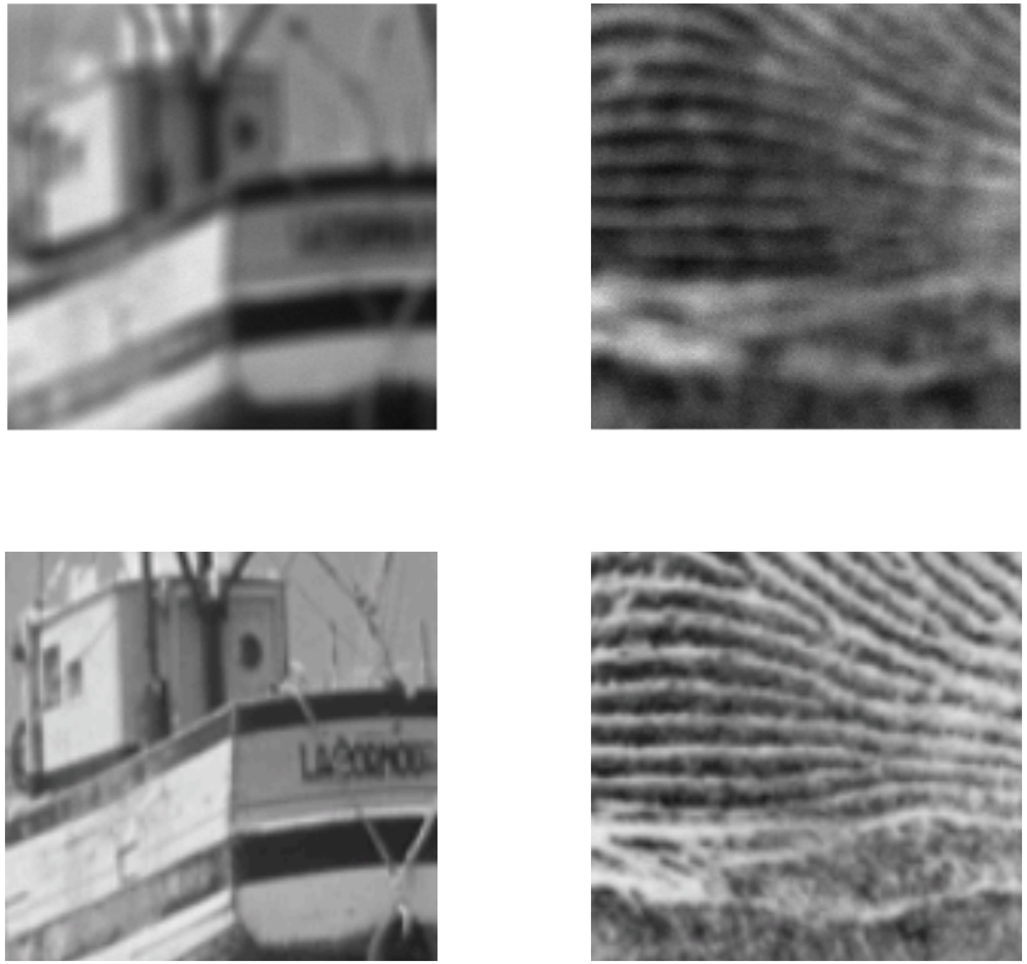
Alireza Chamanzar

Carnegie Mellon University

Electrical and Computer Engineering

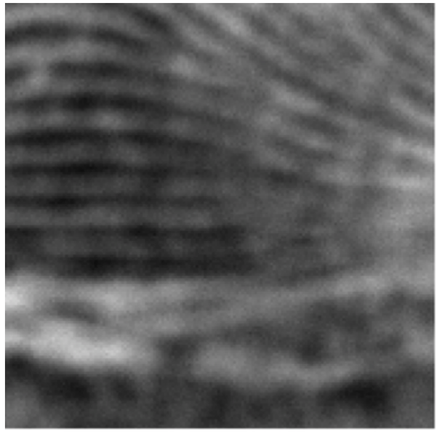
Motivation

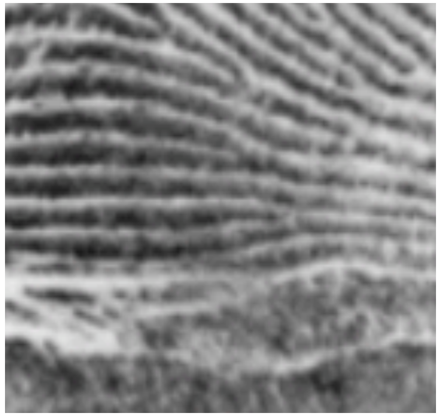
- Image reconstruction\denoising
  - Artifacts (blurring)
  - Noise



Motivation

- Image reconstruction\denoising
  - Artifacts (blurring)
  - Noise





Key idea: Regularization-based restoration methods

Problem formulation

- Rasterized image

$n \times m = N$

y

$N \times 1$

Problem formulation

- Inverse problem

$$y = A_{N\times N} \; f + \boldsymbol{w}$$

$n\times m = N$

$\sim Guassian$

Measured data

Blurring operator

Original data

Problem formulation

- Inverse problem

$$y = A_{N \times N} \, f + \boldsymbol{w}$$

$$\sim Guassian$$

Measured data　　Blurring operator　　Original data

ill-conditioned or non-invertible → Regularization

Regularization

- Additional knowledge\constraints
  - Exploit prior-knowledge

$$J(\mathbf{f}) = \frac{1}{2}\underbrace{\|\mathbf{y} - \mathbf{A}\mathbf{f}\|_2^2}_{\text{Measure of consistency}} + \tau \underbrace{R(\mathbf{f})}_{\text{regularizer}}$$

Regularization

- A good regularizer in image reconstruction?

$$J(\mathbf{f}) = \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{f}\|_2^2 + \tau \boxed{R(\mathbf{f})}$$

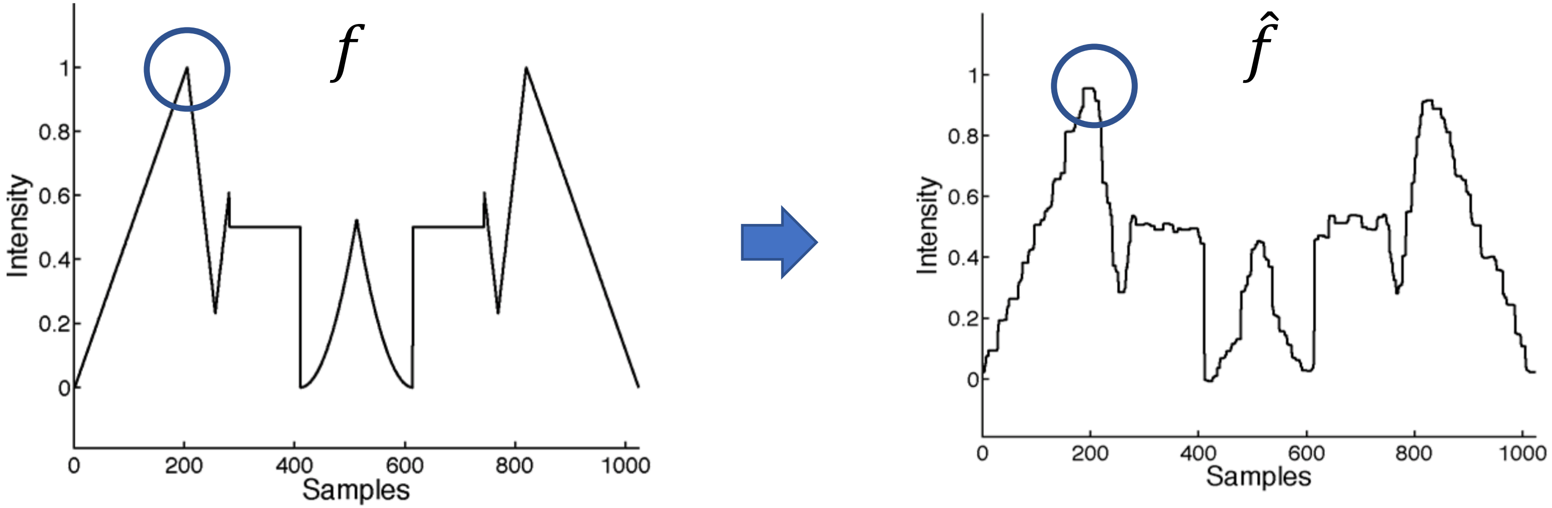
?

Regularization

- A good regularizer in image reconstruction?
  - Preserve **sharpness**

$$J(\mathbf{f}) = \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{f}\|_2^2 + \tau \boxed{R(\mathbf{f})}$$

?

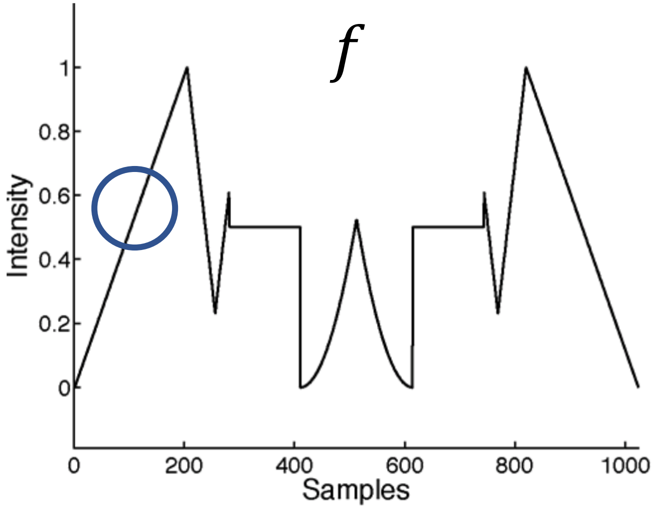


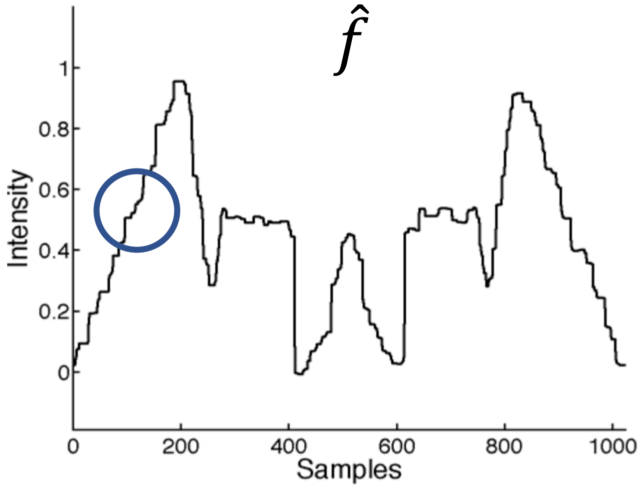
Regularization

- A good regularizer in image reconstruction?
  - Preserve **sharpness**
  - Minimum **staircase** effect

$$J(\mathbf{f}) = \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{f}\|_2^2 + \tau \boxed{R(\mathbf{f})}$$

?



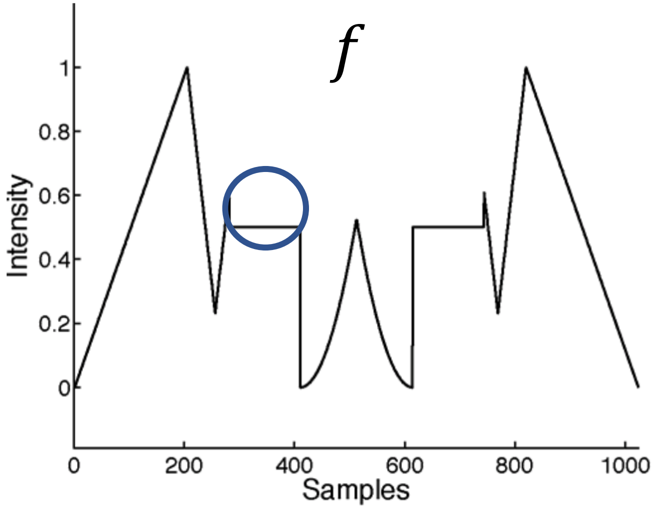


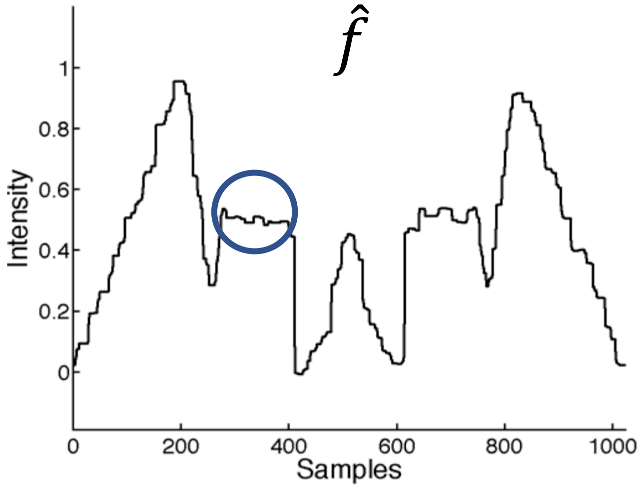
Regularization

- A good regularizer in image reconstruction?
  - Preserve **sharpness**
  - Minimum **staircase** effect
  - Preserve **constant** pieces

$$J(\mathbf{f}) = \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{f}\|_2^2 + \tau \boxed{R(\mathbf{f})}$$

?





Total variance (TV) norm

- A good regularizer in image reconstruction
  - Convex
  - Well-preserved & sharp edges
  - Panelize irregularities in variation

1-D TV

- Definition:

$$\boxed{R(f;p,q) = \int_{\Omega} |D^p f(x)|^q \, dx = \|D^p f\|_q^q}$$

  - where,

f: 1-D signal of finite spatial support $\Omega \subset \mathbb{R}$ with approperiate continuity properties

pth -order derivative : $D^p = \partial^p / \partial x^p$

q-norm: determines how “irregularities” are panelized (**L1** and L2 are widely used, **convex**)

TV: L1 vs L2

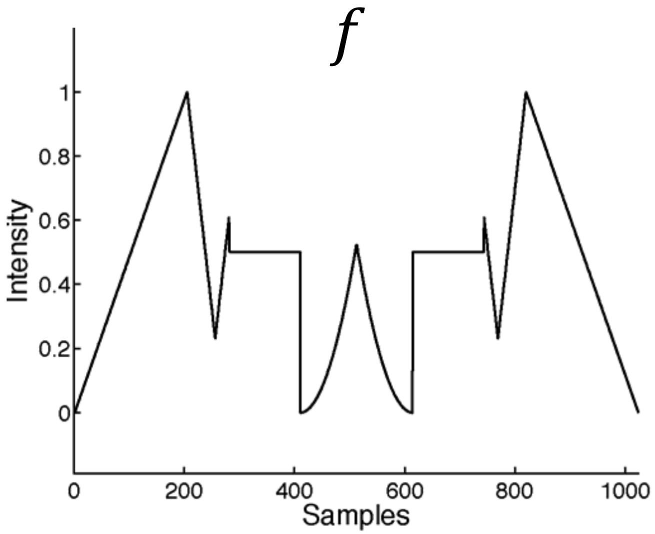
- L1:
  - Less sensitive to **outliers**
  - Better **edge** reconstruction
- L2 (Tikhonov regularization):
  - Closed-form solution
  - Sensitive to outliers

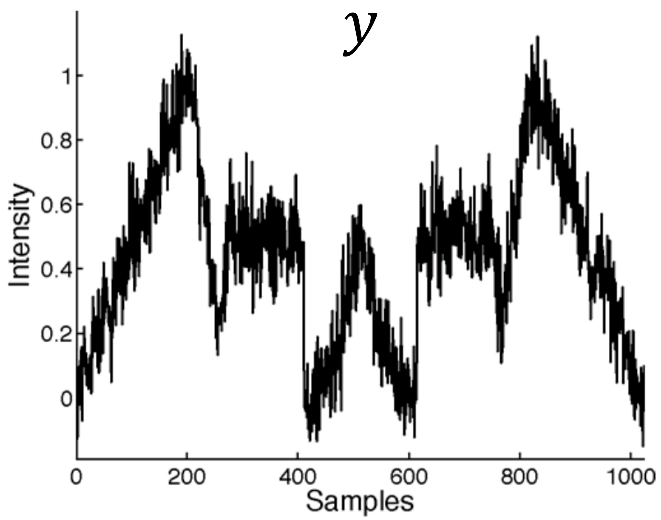
$$\boxed{\hat{f} = arg \min_{f} \frac{1}{2} ||y - Af||_2^2 + \tau ||D^p f||_q^q}$$

TV: L1 vs L2

- Denoising problem:
  - $A = I$ (No blurring, just noise)
  - Gaussian noise ($\sigma = 0.1$)

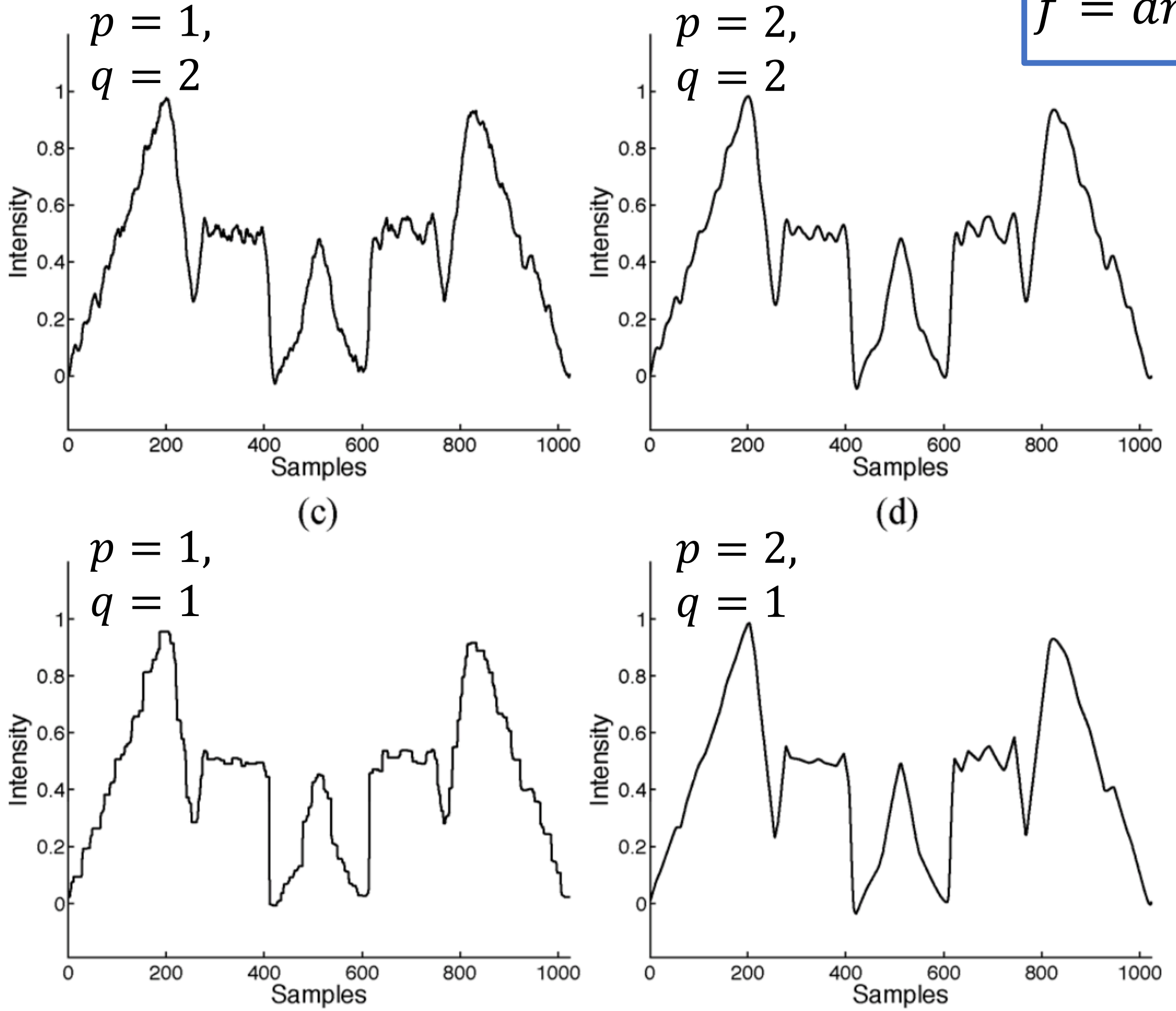
$$\boxed{\hat{f} = arg\min_{f} \frac{1}{2} \|y - Af\|_2^2 + \tau \|D^p f\|_q^q}$$





TV: L1 vs L2

$$\boxed{\hat{f} = arg\min_{f}\frac{1}{2}\|y - f\|_2^2 + \tau\|D^p f\|_q^q}$$



2-D TV

- Definition (**first order**): $$\boxed{\mathrm{TV}(f) = \int_{\boldsymbol{\Omega}} \|\nabla f(\mathbf{x})\|_2 \, d\mathbf{x}.}$$
  - where,

f: 2-D image of continuously differentiable and $\boldsymbol{\Omega} \subset \mathbb{R}^2$.

2-D TV

- Definition (**first order**): $$\boxed{\mathrm{TV}(f) = \int_{\Omega} \|\nabla f(\mathbf{x})\|_2 \, d\mathbf{x}.}$$
  - where,

f: 2-D image of continuously differentiable and $\Omega \subset \mathbb{R}^2$.

Closely related to 1D:

$$\boxed{\mathrm{TV}(f) = \frac{1}{h(q)} \int_{\Omega} \left\| D_\theta^1 f(\mathbf{x}) \right\|_{L_q[0,2\pi]} d\mathbf{x}}$$

  - where,

$$h(q) = \|\cos(\theta)\|_{L_q[0,2\pi]}$$

$$D_\theta^1 f = \langle \nabla f, \mathbf{u}_\theta \rangle = \nabla f \cdot \mathbf{u}_\theta$$

$$\mathbf{u}_\theta = (\cos\theta, \sin\theta)$$

2-D TV

- Definition (**first order**):

$$\boxed{\mathrm{TV}(f) = \int_{\Omega} \|\nabla f(\mathbf{x})\|_2 \, d\mathbf{x}.}$$

  - where,

  f: 2-D image of continuously differentiable and $\Omega \subset \mathbb{R}^2$.

  Closely related to 1D:

$$\boxed{\mathrm{TV}(f) = \frac{1}{h(q)} \int_{\Omega} \left\| D_\theta^1 f(\mathbf{x}) \right\|_{L_q[0,2\pi]} d\mathbf{x}}$$

  - where,

$$h(q) = \|\cos(\theta)\|_{L_q[0,2\pi]}$$

$$D_\theta^1 f = \langle \nabla f, \mathbf{u}_\theta \rangle = \nabla f \cdot \mathbf{u}_\theta$$

$$\mathbf{u}_\theta = (\cos\theta, \sin\theta)$$

equivalents

2-D TV

- Definition (**2nd order**):
  - Different options:
    - Laplacian
    - Frobenius
    - Affine

$$R_L(f) = \int_{\mathbf{\Omega}} |\Delta f(\mathbf{x})|\, d\mathbf{x} \qquad \Delta f(\mathbf{x}) \;=\; f_{xx}(\mathbf{x}) \,+\, f_{yy}(\mathbf{x})$$

$$R_F(f) = \int_{\mathbf{\Omega}} \sqrt{f_{xx}^2(\mathbf{x}) + 2f_{xy}^2(\mathbf{x}) + f_{yy}^2(\mathbf{x})} d\mathbf{x}$$

$$R_A(f) = \int_{\mathbf{\Omega}} \left(\sqrt{f_{xx}^2(\mathbf{x}) + f_{xy}^2(\mathbf{x})} + \sqrt{f_{yx}^2(\mathbf{x}) + f_{yy}^2(\mathbf{x})} \right) d\mathbf{x}.$$

  - All of them are designed mostly for denoising.

2-D TV

- Definition (**2nd order**):
  - This paper's contribution:
    - 2nd order 2D TV for image deblurring (more general than denoising)
    - **Extension to 1st order:**

$$\boxed{\mathrm{TV}(f) = \frac{1}{h(q)} \int_{\mathbf{\Omega}} \left\| D_\theta^1 f(\mathbf{x}) \right\|_{L_q[0,2\pi]} d\mathbf{x}} \Rightarrow \begin{cases} R_S(f) = \displaystyle\int_{\mathbf{\Omega}} \left\| D_{\theta,\phi}^2 f(\mathbf{x}) \right\|_{L_\infty[0,2\pi]^2} d\mathbf{x} \\ R_F(f) = \displaystyle\frac{1}{\pi} \int_{\mathbf{\Omega}} \left\| D_{\theta,\phi}^2 f(\mathbf{x}) \right\|_{L_2[0,2\pi]^2} d\mathbf{x} \end{cases}$$

2-D TV

$$R_S(f) = \int_{\mathbf{\Omega}} \left\| D^2_{\theta,\phi} f(\mathbf{x}) \right\|_{L_\infty[0,2\pi]^2} d\mathbf{x}$$

$$R_F(f) = \frac{1}{\pi} \int_{\mathbf{\Omega}} \left\| D^2_{\theta,\phi} f(\mathbf{x}) \right\|_{L_2[0,2\pi]^2} d\mathbf{x}$$

where,

$$D^2_{\theta,\phi} f(\mathbf{x}) = D^1_\theta (D^1_\phi f) = \mathbf{u}_\theta^T \mathcal{H}_f(\mathbf{x}) \mathbf{v}_\phi$$

Second directional derivative **along θ and ϕ directions**

$$\mathbf{v}_\phi = (\cos\phi, \sin\phi)$$

$$\mathbf{u}_\theta = (\cos\theta, \sin\theta)$$

Directions along θ and ϕ

$$\mathcal{H}_f = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

Hessian Matrix

2-D TV

- Lemma 1:
  - The L_∞ norm of the second directional derivative of f at x is equal to the spectral norm of the Hessian matrix $\left\|\mathcal{H}_f(x)\right\|_2$.
- Proof:
  - The spectral norm of a matrix B is defined as: $\|\mathbf{B}\|_2 = \max_{\|\mathbf{x}\|_2=1} \max_{\|\mathbf{y}\|_2=1} |\mathbf{y}^H\mathbf{B}\mathbf{x}|$

$$\begin{aligned}
\left\|D^2_{\theta,\phi} f(\mathbf{x})\right\|_{L_\infty[0,2\pi]^2} &= \max_{\theta,\phi} \left|D^2_{\theta,\phi} f(\mathbf{x})\right| \\
&= \max_{\theta,\phi} \left|\mathbf{u}_\theta^T \mathcal{H}_f(\mathbf{x})\mathbf{v}_\phi\right| \\
&= \max_{\|\mathbf{u}\|_2=1} \max_{\|\mathbf{v}\|_2=1} \left|\mathbf{u}^T \mathcal{H}_f(\mathbf{x})\mathbf{v}\right| \\
&= \|\mathcal{H}_f(\mathbf{x})\|_2 \, . \; \blacksquare
\end{aligned}$$

2-D TV

- Lemma 2:
  - The L_2 norm of the second directional derivative of f at x is **proportional** to the Frobenius norm of the Hessian matrix $\|\mathcal{H}_f(x)\|_F$.
- Proof:

$\mathcal{H}_f(x)$ is **symmetric** ➡

$$\begin{aligned} D^2_{\theta,\phi} f(\mathbf{x}) &= \mathbf{u}_\theta^T \mathbf{Q} \mathbf{\Lambda}_f(\mathbf{x}) \mathbf{Q}^T \mathbf{v}_\phi \\ &= (\mathbf{Q}^T \mathbf{u}_\theta)^T \mathbf{\Lambda}_f(\mathbf{x}) (\mathbf{Q}^T \mathbf{v}_\phi) \end{aligned}$$

$$\mathbf{Q}^T = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

$$a^2 + b^2 = 1$$

Q is a rotation matrix

➡

$$D^2_{\theta,\phi} f(\mathbf{x}) = \lambda_1 f(\mathbf{x}) \sin(\theta') \sin(\phi') + \lambda_2 f(\mathbf{x}) \cos(\theta') \cos(\phi')$$

2-D TV

- Lemma 2:
  - The L_2 norm of the second directional derivative of f at x is **proportional** to the Frobenius norm of the Hessian matrix $\left\|\mathcal{H}_f(x)\right\|_F$.
- Proof:

$$D^2_{\theta,\phi} f(\mathbf{x}) = \lambda_1 f(\mathbf{x}) \sin(\theta') \sin(\phi') + \lambda_2 f(\mathbf{x}) \cos(\theta') \cos(\phi')$$

$$\Rightarrow \quad \left\|D^2_{\theta,\phi} f(\mathbf{x})\right\|_{L_2[0,2\pi]^2} = \left(\int_0^{2\pi} \int_0^{2\pi} \left|D^2_{\theta,\phi} f(\mathbf{x})\right|^2 d\theta d\phi \right)^{1/2}$$

$$= \sqrt{\pi^2 \left[\lambda_1^2 f(\mathbf{x}) + \lambda_2^2 f(\mathbf{x})\right]}$$

$$= \pi \left\|\mathcal{H}_f(\mathbf{x})\right\|_F . \blacksquare$$

Spectral-norm Regularizer

- Interpretation:

$$\boxed{R_S(f) = \int_{\Omega} \left\| D^2_{\theta,\phi} f(\mathbf{x}) \right\|_{L_\infty [0,2\pi]^2} d\mathbf{x}}$$

  - Alternative Def. of spectral norm:

$$\|\mathcal{H}_f(\mathbf{x})\|_2 = \max_{i=1,2} (|\lambda_i f(\mathbf{x})|)$$

where,

$$\lambda_{1,2} f(\mathbf{x}) = \frac{\Delta f(\mathbf{x}) \pm \sqrt{\left(\bar{\Delta} f(\mathbf{x})\right)^2 + \left(\Gamma f(\mathbf{x})\right)^2}}{2}$$

Scalar Operators	Vectorial Operators
$\Delta = \partial_{xx} + \partial_{yy}$	$\nabla = (\partial_x, \partial_y)$
$\bar{\Delta} = \partial_{xx} - \partial_{yy}$	$\mathcal{U} = (\partial_{xx} - \partial_{yy}, 2\partial_{xy})$
$\Gamma = 2\partial_{xy}$	$\mathcal{V} = \left(\partial_{xx}, \sqrt{2}\partial_{xy}, \partial_{yy}\right)$

(Lemma 1) $\left\| D^2_{\theta,\phi} f(\mathbf{x}) \right\|_{L_\infty [0,2\pi]^2} = \|\mathcal{H}_f(\mathbf{x})\|_2$

$$\max (|\alpha + \beta|, |\alpha - \beta|) = |\alpha| + \beta, \quad \forall \beta \geq 0$$

$$\boxed{R_S(f) = \frac{1}{2} \int_{\Omega} (|\Delta f(\mathbf{x})| + \|\mathcal{U} f(\mathbf{x})\|_2)\, d\mathbf{x}}$$

Discrete Problem formulation

- Discretize:
  - FW finite differences
  - Periodic boundary conditions

$$\boxed{J(\mathbf{f}) = \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{f}\|_2^2 + \tau R(\mathbf{f})}$$

$$\begin{aligned}
\mathbf{f}_{xx}[i,j] &= \mathbf{f}[i,j] - 2\mathbf{f}[i+1,j] + \mathbf{f}[i+2,j],\\
\mathbf{f}_{yy}[i,j] &= \mathbf{f}[i,j] - 2\mathbf{f}[i,j+1] + \mathbf{f}[i,j+2],\\
\mathbf{f}_{xy}[i,j] &= \mathbf{f}[i,j] - \mathbf{f}[i+1,j] - \mathbf{f}[i,j+1] + \mathbf{f}[i+1,j+1].
\end{aligned}$$

Discrete Problem formulation

- Discretize:
  - FW finite differences
  - Periodic boundary conditions

Scalar Operators	Vectorial Operators
$\Delta = \partial_{xx} + \partial_{yy}$	$\nabla = (\partial_x, \partial_y)$
$\bar{\Delta} = \partial_{xx} - \partial_{yy}$	$\mathcal{U} = (\partial_{xx} - \partial_{yy}, 2\partial_{xy})$
$\Gamma = 2\partial_{xy}$	$\mathcal{V} = \left(\partial_{xx}, \sqrt{2}\partial_{xy}, \partial_{yy}\right)$

$$R_L(f) = \int_{\mathbf{\Omega}} |\Delta f(\mathbf{x})| \, d\mathbf{x}$$

$$R_F(f) = \int_{\mathbf{\Omega}} \sqrt{f_{xx}^2(\mathbf{x}) + 2f_{xy}^2(\mathbf{x}) + f_{yy}^2(\mathbf{x})} d\mathbf{x}$$

$$R_S(f) = \frac{1}{2} \int_{\mathbf{\Omega}} \left(|\Delta f(\mathbf{x})| + \|\mathcal{U} f(\mathbf{x})\|_2\right) d\mathbf{x}$$

$$R_L(\mathbf{f}) = \|\mathbf{\Delta f}\|_1 = \sum_{i=1}^{N} |(\mathbf{\Delta f})_i|$$

$$R_F(\mathbf{f}) = \|\mathbf{Vf}\|_1 = \frac{\sqrt{2}}{2} \sum_{i=1}^{N} \sqrt{(\mathbf{\Delta f})_i^2 + (\bar{\mathbf{\Delta}}\mathbf{f})_i^2 + (\mathbf{\Gamma f})_i^2}$$

$$R_S(\mathbf{f}) = \frac{1}{2}\left(\|\mathbf{\Delta f}\|_1 + \|\mathbf{Uf}\|_1\right) = \frac{1}{2}\left(\sum_{i=1}^{N} |(\mathbf{\Delta f})_i| + \sqrt{(\bar{\mathbf{\Delta}}\mathbf{f})_i^2 + (\mathbf{\Gamma f})_i^2}\right)$$

Discrete Problem formulation

$$J(\mathbf{f}) = \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{f}\|_2^2 + \tau R(\mathbf{f})$$

- Majorization minimization (MM):
  - Iterative algorithm

$$\mathbf{f}^* = \arg\min_{\mathbf{f}} J(\mathbf{f}) \quad \Rightarrow \quad \mathbf{f}^{(t+1)} = \arg\min_{\mathbf{f}} Q\left(\mathbf{f}; \mathbf{f}^{(t)}\right)$$

where, Q is majorizer of f as follow:

$$\left.\begin{aligned} Q\left(\mathbf{f}; \mathbf{f}^{(t)}\right) &\geq J(\mathbf{f}), \quad \forall \mathbf{f} \\ Q\left(\mathbf{f}^{(t)}; \mathbf{f}^{(t)}\right) &= J\left(\mathbf{f}^{(t)}\right). \end{aligned}\right\} \Rightarrow$$

Q is minimized => f is minimized

Majorization minimization (MM)

- Finding majorizers for $\boldsymbol{R_L, R_F, and\ R_S}$

$$\boxed{Q\left(\mathbf{f};\mathbf{f}^{(t)}\right) \geq J(\mathbf{f}),}$$

$$|g(x)| \leq \frac{|g(y)|}{2} + \frac{g(x)^2}{2\,|g(y)|}, \quad \forall x, \forall y : g(y) \neq 0$$

$$R_L(\mathbf{f}) = \|\boldsymbol{\Delta}\mathbf{f}\|_1 = \sum_{i=1}^{N} |(\boldsymbol{\Delta}\mathbf{f})_i|$$

$$\Rightarrow \boxed{Q_L\left(\mathbf{f};\mathbf{f}^{(t)}\right) = \frac{1}{2}\left\|\boldsymbol{\Delta}\mathbf{f}^{(t)}\right\|_1 + \frac{1}{2}\sum_{i=1}^{N}\frac{(\boldsymbol{\Delta}\mathbf{f})_i^2}{\left|\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i\right|}}$$

$$\sqrt{g(x)} \leq \frac{\sqrt{g(y)}}{2} + \frac{g(x)}{2\sqrt{g(x)}}, \quad \forall x : g(x) > 0, \forall y : g(y) \geq 0$$

$$R_F(\mathbf{f}) = \|\mathbf{V}\mathbf{f}\|_1 = \frac{\sqrt{2}}{2}\sum_{i=1}^{N}\sqrt{(\boldsymbol{\Delta}\mathbf{f})_i^2 + (\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2}$$

$$\Rightarrow \boxed{Q_F\left(\mathbf{f};\mathbf{f}^{(t)}\right) = \frac{1}{2}\left\|\mathbf{V}\mathbf{f}^{(t)}\right\|_1 + \frac{\sqrt{2}}{4} \times \sum_{i=1}^{N}\frac{\left[(\boldsymbol{\Delta}\mathbf{f})_i^2 + (\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2\right]}{\sqrt{\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i^2 + \left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}}}$$

Majorization minimization (MM)

- Finding majorizers for $\boldsymbol{R_L, R_F, and\ R_S}$

$$\sqrt{g(x)} \leq \frac{\sqrt{g(y)}}{2} + \frac{g(x)}{2\sqrt{g(x)}}, \quad \forall x : g(x) > 0, \forall y : g(y) \geq 0$$

$$R_S(\mathbf{f}) = \frac{1}{2}\left(\|\boldsymbol{\Delta}\mathbf{f}\|_1 + \|\mathbf{U}\mathbf{f}\|_1\right)$$

$$= \frac{1}{2}\left(\sum_{i=1}^{N} |(\boldsymbol{\Delta}\mathbf{f})_i| + \sqrt{(\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2}\right)$$

$$Q_L\left(\mathbf{f};\mathbf{f}^{(t)}\right) = \frac{1}{2}\left\|\boldsymbol{\Delta}\mathbf{f}^{(t)}\right\|_1 + \frac{1}{2}\sum_{i=1}^{N} \frac{(\boldsymbol{\Delta}\mathbf{f})_i^2}{\left|\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i\right|}$$

$$Q_F\left(\mathbf{f};\mathbf{f}^{(t)}\right) = \frac{1}{2}\left\|\mathbf{V}\mathbf{f}^{(t)}\right\|_1 + \frac{\sqrt{2}}{4} \times \sum_{i=1}^{N} \frac{\left[(\boldsymbol{\Delta}\mathbf{f})_i^2 + (\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2\right]}{\sqrt{\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i^2 + \left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}}$$

$$Q_U\left(\mathbf{f};\mathbf{f}^{(t)}\right) = \frac{1}{2}\left\|\mathbf{U}\mathbf{f}^{(t)}\right\|_1 + \frac{1}{2}\sum_{i=1}^{N} \frac{\left[(\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2\right]}{\sqrt{\left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}}$$

Majorization minimization (MM)

- Finding majorizers for $\boldsymbol{R_L, R_F, and\ R_S}$

$$Q_L\left(\mathbf{f};\mathbf{f}^{(t)}\right) = \frac{1}{2}\left\|\boldsymbol{\Delta}\mathbf{f}^{(t)}\right\|_1 + \frac{1}{2}\sum_{i=1}^{N}\frac{(\boldsymbol{\Delta}\mathbf{f})_i^2}{\left|\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i\right|}$$

$$Q_F\left(\mathbf{f};\mathbf{f}^{(t)}\right) = \frac{1}{2}\left\|\mathbf{V}\mathbf{f}^{(t)}\right\|_1 + \frac{\sqrt{2}}{4} \times \sum_{i=1}^{N}\frac{\left[(\boldsymbol{\Delta}\mathbf{f})_i^2 + (\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2\right]}{\sqrt{\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i^2 + \left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}}$$

$$Q_U\left(\mathbf{f};\mathbf{f}^{(t)}\right) = \frac{1}{2}\left\|\mathbf{U}\mathbf{f}^{(t)}\right\|_1 + \frac{1}{2}\sum_{i=1}^{N}\frac{\left[(\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2\right]}{\sqrt{\left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}}$$

$$\begin{aligned} R_S(\mathbf{f}) &= \frac{1}{2}\left(\|\boldsymbol{\Delta}\mathbf{f}\|_1 + \|\mathbf{U}\mathbf{f}\|_1\right) \\ &= \frac{1}{2}\left(\sum_{i=1}^{N}|(\boldsymbol{\Delta}\mathbf{f})_i| + \sqrt{(\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2}\right) \end{aligned}$$

$$\Rightarrow$$

$$\begin{aligned} Q_S\left(\mathbf{f};\mathbf{f}^{(t)}\right) &= \frac{1}{2}\left(Q_L\left(\mathbf{f};\mathbf{f}^{(t)}\right) + Q_U\left(\mathbf{f};\mathbf{f}^{(t)}\right)\right) \\ &= \frac{1}{4}\sum_{i=1}^{N}\frac{\left[(\bar{\boldsymbol{\Delta}}\mathbf{f})_i^2 + (\boldsymbol{\Gamma}\mathbf{f})_i^2\right]}{\sqrt{\left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}} \\ &\quad + \frac{1}{4}\sum_{i=1}^{N}\frac{(\boldsymbol{\Delta}\mathbf{f})_i^2}{\left|\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i\right|} + \text{const.} \end{aligned}$$

Majorization minimization (MM)

- Regularizer majorizer (Q_R) **compact form**:

$$Q_R\left(\mathbf{f};\mathbf{f}^{(t)}\right) = \frac{1}{2}\mathbf{f}^T\left(\mathbf{M}^T\mathbf{W}^{(t)}\mathbf{M}\right)\mathbf{f} + \text{const.}$$

where,

Q_L $\quad\Rightarrow\quad$ $\mathbf{M} = \boldsymbol{\Delta}$

Q_F, Q_S $\quad\Rightarrow\quad$ $\mathbf{M} = [\boldsymbol{\Delta}^T, \bar{\boldsymbol{\Delta}}^T, \boldsymbol{\Gamma}^T]^T$

$$\mathbf{W}^{(t)} = \begin{bmatrix} \mathbf{W}_1^{(t)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_2^{(t)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{W}_2^{(t)} \end{bmatrix}_{3N\times N} \Rightarrow \left\{ \begin{aligned} \left(\mathbf{W}_L^{(t)}\right)_{ii} &= \frac{1}{2\left|\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i\right|}, \quad i = 1\ldots N \\ \left(\mathbf{W}_U^{(t)}\right)_{ii} &= \frac{1}{2\sqrt{\left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}} \\ \left(\mathbf{W}_F^{(t)}\right)_{ii} &= \frac{\sqrt{2}}{2\sqrt{\left(\boldsymbol{\Delta}\mathbf{f}^{(t)}\right)_i^2 + \left(\bar{\boldsymbol{\Delta}}\mathbf{f}^{(t)}\right)_i^2 + \left(\boldsymbol{\Gamma}\mathbf{f}^{(t)}\right)_i^2}}. \end{aligned} \right.$$

Spectral

$$\boxed{\mathbf{W}_1^{(t)} = \mathbf{W}_L^{(t)} \text{ and } \mathbf{W}_2^{(t)} = \mathbf{W}_U^{(t)}}$$

Frobenius

$$\boxed{\mathbf{W}_1^{(t)} = \mathbf{W}_2^{(t)} = \mathbf{W}_F^{(t)}}$$

Majorization minimization (MM)

- Quadratic majorizer for objective function (J):

$$J(\mathbf{f}) = \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{f}\|_2^2 + \tau R(\mathbf{f}) \quad \Rightarrow \quad \begin{aligned} Q\left(\mathbf{f};\mathbf{f}^{(t)}\right) &= \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{f}\|_2^2 + \tau Q_R\left(\mathbf{f};\mathbf{f}^{(t)}\right) \\ &= \frac{1}{2}\mathbf{f}^T\left(\mathbf{A}^T\mathbf{A} + \tau\mathbf{M}^T\mathbf{W}^{(t)}\mathbf{M}\right)\mathbf{f} \\ &\quad - 2\mathbf{f}^T\mathbf{A}^T\mathbf{y} + \text{const.} \end{aligned}$$

$$\frac{dQ}{df} = 0 \quad \Rightarrow \quad \boxed{\underbrace{\left(\mathbf{A}^T\mathbf{A} + \tau\mathbf{M}^T\mathbf{W}^{(t)}\mathbf{M}\right)}_{\mathbf{S}^{(t)}}\mathbf{f}^{(t+1)} = \mathbf{A}^T\mathbf{y}.} \quad \Rightarrow$$

Solve using conjugate gradient method (CG)

Inversion of very large matrix S!!

Conjugate gradient (CG)

- Conjugate gradient (CG):

$$\boxed{\underbrace{\left(\mathbf{A}^T\mathbf{A}+\tau\mathbf{M}^T\mathbf{W}^{(t)}\mathbf{M}\right)}_{\mathbf{S}^{(t)}}\mathbf{f}^{(t+1)}=\mathbf{A}^T\mathbf{y}.}$$

$$\boldsymbol{Ax}=\boldsymbol{b}$$

$$\mathbf{u}^{\mathsf{T}}\mathbf{A}\mathbf{v}=0.$$

$$\langle\mathbf{u},\mathbf{v}\rangle_{\mathbf{A}}:=\langle\mathbf{A}\mathbf{u},\mathbf{v}\rangle=\langle\mathbf{u},\mathbf{A}^{\mathsf{T}}\mathbf{v}\rangle=\langle\mathbf{u},\mathbf{A}\mathbf{v}\rangle=\mathbf{u}^{\mathsf{T}}\mathbf{A}\mathbf{v}.$$

$$P=\{\mathbf{p}_1,\dots,\mathbf{p}_n\}\ \Rightarrow\ \mathbf{x}_*=\sum_{i=1}^{n}\alpha_i\mathbf{p}_i$$

https://en.wikipedia.org/wiki/Conjugate_gradient_method

Conjugate gradient (CG)

- Conjugate gradient (CG):

$$\boxed{\underbrace{\left(\mathbf{A}^T\mathbf{A} + \tau\mathbf{M}^T\mathbf{W}^{(t)}\mathbf{M}\right)}_{\mathbf{S}^{(t)}}\mathbf{f}^{(t+1)} = \mathbf{A}^T\mathbf{y}.}$$

$$\boldsymbol{Ax = b}$$

$$\mathbf{x}_* = \sum_{i=1}^{n}\alpha_i\mathbf{p}_i$$

$$\mathbf{A}\mathbf{x}_* = \sum_{i=1}^{n}\alpha_i\mathbf{A}\mathbf{p}_i$$

$$\mathbf{p}_k^\mathsf{T}\mathbf{A}\mathbf{x}_* = \sum_{i=1}^{n}\alpha_i\mathbf{p}_k^\mathsf{T}\mathbf{A}\mathbf{p}_i \qquad \text{(Multiply left by } \mathbf{p}_k^\mathsf{T}\text{)}$$

$$\mathbf{p}_k^\mathsf{T}\mathbf{b} = \sum_{i=1}^{n}\alpha_i\langle\mathbf{p}_k,\mathbf{p}_i\rangle_\mathbf{A} \qquad (\mathbf{A}\mathbf{x}_* = \mathbf{b} \text{ and } \langle\mathbf{u},\mathbf{v}\rangle_\mathbf{A} = \mathbf{u}^\mathsf{T}\mathbf{A}\mathbf{v})$$

$$\langle\mathbf{p}_k,\mathbf{b}\rangle = \alpha_k\langle\mathbf{p}_k,\mathbf{p}_k\rangle_\mathbf{A} \qquad (\mathbf{u}^\mathsf{T}\mathbf{v} = \langle\mathbf{u},\mathbf{v}\rangle \text{ and } \forall i\neq k: \langle\mathbf{p}_k,\mathbf{p}_i\rangle_\mathbf{A} = 0) \Rightarrow \boxed{\alpha_k = \frac{\langle\mathbf{p}_k,\mathbf{b}\rangle}{\langle\mathbf{p}_k,\mathbf{p}_k\rangle_\mathbf{A}}}$$

https://en.wikipedia.org/wiki/Conjugate_gradient_method

Conjugate gradient (CG)

- Conjugate gradient (CG):

$$\boxed{\alpha_k = \frac{\langle \mathbf{p}_k, \mathbf{b} \rangle}{\langle \mathbf{p}_k, \mathbf{p}_k \rangle_{\mathbf{A}}}}$$

$$\boxed{f(\mathbf{x}) = \tfrac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} - \mathbf{x}^\mathsf{T}\mathbf{b}}$$

$$\boxed{\underbrace{\left(\mathbf{A}^T\mathbf{A} + \tau\mathbf{M}^T\mathbf{W}^{(t)}\mathbf{M}\right)}_{\mathbf{S}^{(t)}} \mathbf{f}^{(t+1)} = \mathbf{A}^T\mathbf{y}.}$$

$$\boldsymbol{Ax = b}$$

$$\mathbf{r}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_k.$$

$$\mathbf{p}_k = \mathbf{r}_k - \sum_{i<k} \frac{\mathbf{p}_i^\mathsf{T}\mathbf{A}\mathbf{r}_k}{\mathbf{p}_i^\mathsf{T}\mathbf{A}\mathbf{p}_i}\mathbf{p}_i$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$$

$$\Rightarrow \quad \alpha_k = \frac{\mathbf{p}_k^\mathsf{T}\mathbf{b}}{\mathbf{p}_k^\mathsf{T}\mathbf{A}\mathbf{p}_k} = \frac{\mathbf{p}_k^\mathsf{T}(\mathbf{r}_k + \mathbf{A}\mathbf{x}_k)}{\mathbf{p}_k^\mathsf{T}\mathbf{A}\mathbf{p}_k} = \frac{\mathbf{p}_k^\mathsf{T}\mathbf{r}_k}{\mathbf{p}_k^\mathsf{T}\mathbf{A}\mathbf{p}_k},$$

https://en.wikipedia.org/wiki/Conjugate_gradient_method

Conjugate gradient (CG)

- Conjugate gradient (CG):

$$\underbrace{\left(\mathbf{A}^T\mathbf{A} + \tau\mathbf{M}^T\mathbf{W}^{(t)}\mathbf{M}\right)}_{\mathbf{S}^{(t)}}\mathbf{f}^{(t+1)} = \mathbf{A}^T\mathbf{y}.$$

$$\boldsymbol{Ax = b}$$

$\mathbf{r}_0 := \mathbf{b} - \mathbf{A}\mathbf{x}_0$
$\mathbf{p}_0 := \mathbf{r}_0$
$k := 0$
repeat

$$\alpha_k := \frac{\mathbf{r}_k^\mathsf{T}\mathbf{r}_k}{\mathbf{p}_k^\mathsf{T}\mathbf{A}\mathbf{p}_k}$$

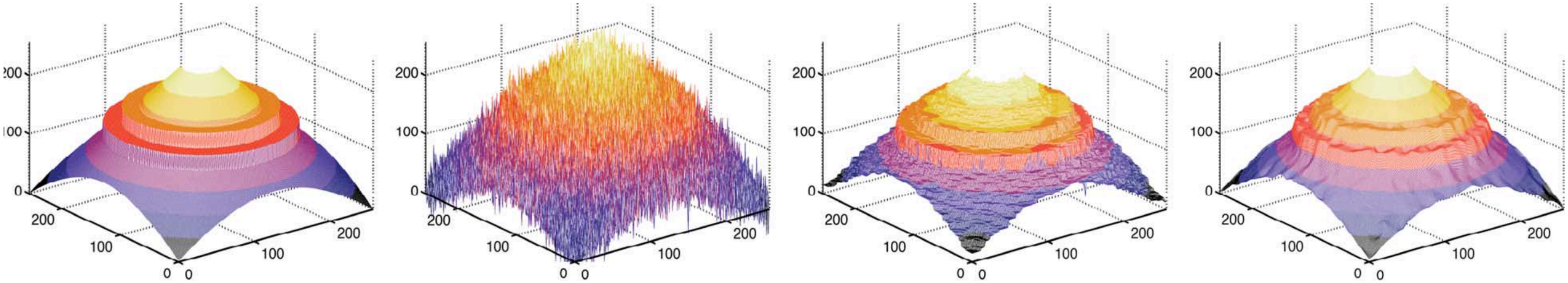
$\mathbf{x}_{k+1} := \mathbf{x}_k + \alpha_k\mathbf{p}_k$
$\mathbf{r}_{k+1} := \mathbf{r}_k - \alpha_k\mathbf{A}\mathbf{p}_k$
if r_{k+1} is sufficiently small, then exit loop

$$\beta_k := \frac{\mathbf{r}_{k+1}^\mathsf{T}\mathbf{r}_{k+1}}{\mathbf{r}_k^\mathsf{T}\mathbf{r}_k}$$

$\mathbf{p}_{k+1} := \mathbf{r}_{k+1} + \beta_k\mathbf{p}_k$
$k := k + 1$
end repeat
The result is $\mathbf{x}_{k+1}$

https://en.wikipedia.org/wiki/Conjugate_gradient_method

Results



(a) (b) (c) (d)

f $y, \sigma = 20$ $\hat{f}_{TV}$ $\hat{f}_{Rs}$