Tdsi AR
21/01/2021
Durée : 1 heure (QCM compris)

NOM : ______________________
Prénom : ______________________

Cet énoncé est composé de 7 pages (y compris celle-ci). Merci de compléter vos nom et prénom en haut à droite de la première page et de placer vos initiales sur les pages suivantes.

Tous les documents sont autorisés.

Exercice (14 points)

In this exercise, we solve the computed tomography reconstruction problem by minimizing the squared error

$$\mathcal{C}(\mathbf{f}) = \|\mathbf{s} - \mathbf{R}\mathbf{f}\|_2^2 + \alpha \mathcal{J}(\mathbf{f}) \tag{1}$$

where $\mathbf{s} \in \mathbb{R}^{M\times 1}$ represents the measurement vector, $\mathbf{R} \in \mathbb{R}^{M\times N}$ the discrete Radon transform, $\mathbf{f} \in \mathbb{R}^{N\times 1}$ the attenuation (unknown) image, α is the regularization parameter and $\mathcal{J}$ a regularization term. As depicted in figure 1, we will consider two types of regularizers $\mathcal{J}$: the $\ell 2$-norm and the $\ell 1$-norm of the gradient of the (unknown) image. The latter is referred to as the total variation. To minimize Eq. (1), we implement a simple gradient descent with the update rule

$$\mathbf{f}^{(k)} = \mathbf{f}^{(k-1)} - \tau \mathbf{g}^{(k)}, \quad \mathbf{f}^{(0)} = \mathbf{0} \tag{2}$$

where τ is the step length and $\mathbf{g}^{(k)}$ is the gradient of the cost function $\mathcal{C}$.

Notations and definitions We define the $\ell 2$-norm of the vector $\mathbf{f} \in \mathbb{R}^N$ as $\|\mathbf{f}\|_2^2 = \sum_{n=1}^{N} f_n^2$ and the $\ell 1$-norm as $\|\mathbf{f}\|_1 = \sum_n |f_n|$. We denote the n-th entry of the vector $\mathbf{f}$ as $(\mathbf{f})_n$ or f_n. We denote $(\mathbf{F})_{n,m}$ or $F_{n,m}$ the entry of the n-th row and m-th column of $\mathbf{F}$.

We recall the definition of the gradient of a function $\mathcal{C} : \mathbb{R}^N \to \mathbb{R}$

$$\nabla \mathcal{C} = \begin{bmatrix} \frac{\partial}{\partial f_1}\mathcal{C} \\ \vdots \\ \frac{\partial}{\partial f_N}\mathcal{C} \end{bmatrix}$$

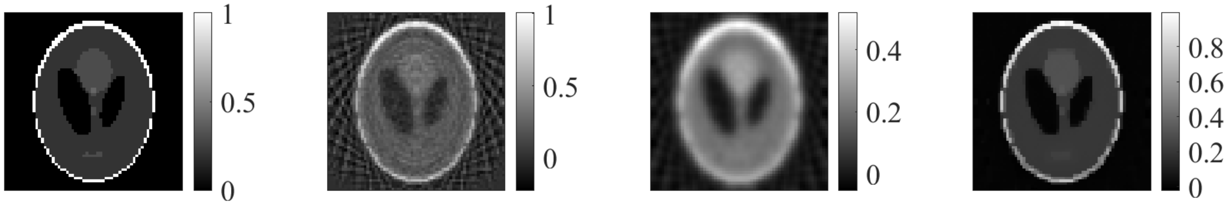

Figure 1 – From left to right : original image, reconstruction with no regularization (α=0), using an ℓ-2 norm regularization term, using an ℓ-1 norm regularization term (total variation).

Questions

1. Problem statement

   (a) Here, we consider the limited angle computed tomography problem for which we have 18 view angles over 180. Assuming the detector has 95 pixels and the image is 64×64, specify the dimensions N and M defined by Eq. (1).

   > **Solution:**
   >
   > $$M = 95 \times 18 = 1710$$
   > $$N = 64 \times 64 = 4096$$

   (b) Why is this problem referred to as "limited angle"?

   > **Solution:** We have more unknowns than measurements

   (c) Let $\mathbf{F} \in \mathbb{R}^{N_x \times N_y}$ be the unknown image as depicted at figure 1. Give the relationship between the matrix $\mathbf{F}$ and the vector $\mathbf{f}$ defined in Eq. (1), for $N_x = N_y = 3$.

   > **Solution:** $\mathbf{x}$ is obtained from $\mathbf{X}$ through vectorization
   >
   > $$\mathbf{x} = [X_{11}, X_{21}, X_{31}, X_{12}, X_{22}, X_{32}, X_{13}, X_{23}, X_{33}]^\top$$

2. We start by investigating the case $\alpha = 0$, i.e., $\mathcal{C}(\mathbf{f}) = \|\mathbf{s} - \mathbf{R}\mathbf{f}\|_2^2$ that we rewrite to derive its gradient.

   (a) Suppose $\mathbf{a} \in \mathbb{R}^N$. Show that $\|\mathbf{a}\|^2 = \mathbf{a}^\top \mathbf{a}$.

   > **Solution:** By definition of the ℓ_2-norm
   >
   > $$\|\mathbf{a}\|^2 = \sum_k a_k^2$$

   (b) Suppose $\mathbf{a} \in \mathbb{R}^N$ and $\mathbf{b} \in \mathbb{R}^N$. Show that $\mathbf{a}^\top \mathbf{b} = \mathbf{b}^\top \mathbf{a}$.

   > **Solution:** As any scalar $a \in \mathbb{R}$ is such that $a^\top = a$. Therefore
   >
   > $$\mathbf{a}^\top \mathbf{b} = \left(\mathbf{a}^\top \mathbf{b}\right)^\top = \mathbf{b}^\top \mathbf{a}.$$

   (c) Show that $\mathcal{C}(\mathbf{f}) = \mathbf{s}^\top \mathbf{s} - 2\mathbf{s}^\top \mathbf{R}\mathbf{f} + \mathbf{f}^\top \mathbf{R}^\top \mathbf{R}\mathbf{f}$.

Solution:

$$\begin{aligned}\|\mathbf{s}-\mathbf{Rf}\|^2 &= (\mathbf{s}-\mathbf{Rf})^\top(\mathbf{s}-\mathbf{Rf}) = \mathbf{s}^\top\mathbf{s} - (\mathbf{Rf})^\top\mathbf{s} - \mathbf{s}^\top(\mathbf{Rf}) + (\mathbf{Rf})^\top\mathbf{Rf}\\ &= \mathbf{s}^\top\mathbf{s} - 2\mathbf{s}^\top(\mathbf{Rf}) + (\mathbf{Rf})^\top\mathbf{Rf}\end{aligned}$$

3. We now compute the gradient of $\mathcal{C}(\mathbf{f}) = \|\mathbf{s}-\mathbf{Rf}\|_2^2$.
   (a) Suppose some constant $\mathbf{a} \in \mathbb{R}^N$. Show that the gradient of $\mathcal{D}(\mathbf{f}) = \mathbf{a}^\top\mathbf{a}$ is given by $\nabla\mathcal{D} = \mathbf{0}$.

Solution: Computing the i-th component of the gradient

$$\frac{\partial}{\partial f_n}\mathcal{D} = \frac{\partial}{\partial f_n}\sum_k a_k^2 = 0, \quad \forall i.$$

Therefore, $\nabla\mathcal{D} = \mathbf{0}$.

   (b) Suppose $\mathbf{a} \in \mathbb{R}^N$. Show that the gradient of $\mathcal{D}(\mathbf{f}) = \mathbf{a}^\top\mathbf{f}$ is given by $\nabla\mathcal{D} = \mathbf{a}$.

Solution: Computing the i-th component of the gradient

$$\frac{\partial}{\partial f_n}\mathcal{D} = \frac{\partial}{\partial f_n}\sum_k a_k f_k = a_n, \quad \forall n.$$

Therefore, $\nabla\mathcal{D} = \mathbf{a}$.

   (c) Suppose $\mathbf{L} \in \mathbb{R}^{M\times N}$. Show that the gradient of $\mathcal{D}(\mathbf{f}) = \mathbf{f}^\top\mathbf{L}^\top\mathbf{Lf}$ is given by $\nabla\mathcal{D} = 2\mathbf{L}^\top\mathbf{Lx}$.

Solution: Computing the i-th component of the gradient

$$\begin{aligned}\frac{\partial}{\partial f_n}\mathcal{D} &= \frac{\partial}{\partial f_n}\sum_k (\mathbf{Lx})_k^2 = \sum_k \frac{\partial}{\partial f_n}(\mathbf{Lx})_k^2\\ &= \sum_k \frac{\partial}{\partial f_n}\left(\sum_\ell L_{k,\ell}\, f_\ell\right)^2 = \sum_k 2L_{k,n}\left(\sum_\ell L_{k,\ell}\, f_\ell\right)\\ &= \sum_k 2L_{k,n}\,(\mathbf{Lf})_k\\ &= 2\sum_k L_{n,k}^\top\,(\mathbf{Lf})_k = 2\left(\mathbf{L}^\top\mathbf{Lf}\right)_n\end{aligned}$$

Therefore, $\nabla\mathcal{D} = 2\mathbf{L}^\top\mathbf{Lf}$.

(d) Derive the gradient of the data fidelity term $\mathcal{C}(\mathbf{f}) = \|\mathbf{s} - \mathbf{R}\mathbf{f}\|_2^2$.

Solution: By linearity, the previous results leads to

$$\nabla \mathcal{C} = 2\mathbf{R}^\top \mathbf{s} + 2\mathbf{R}^\top \mathbf{R}\mathbf{f}$$

4. We will consider regularizers that rely on the image gradient $\nabla \mathbf{f}$, where ∇ is an operator that computes the gradient of an image. Specifically, we define

$$\nabla = \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \end{bmatrix},$$

where $\mathbf{D}_1 \in \mathbb{R}^N$ computes the discrete gradient along the first dimension and $\mathbf{D}_2 \in \mathbb{R}^N$ along the second dimension. Note that this operator differs from the gradient that applies to real-valued functions.

(a) What is the dimension of ∇ ? The dimension of $\nabla \mathbf{f}$?

Solution: $\nabla \in \mathbb{R}^{2N \times N}$, $\nabla \mathbf{f} \in \mathbb{R}^{2N}$

(b) Assuming $(\mathbf{D}_1 \mathbf{f})_{n,m} = F_{n+1,m} - F_{n,m}$, give the expression of the horizontal gradient operator $\mathbf{D}_1$ when $N = 3 \times 3$ (all matrix coefficients are expected).

Solution:

$$\mathbf{D}_1 = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

(c) How is the gradient computed on the border of the image ? What is the underlying assumption ?

Solution: $(\mathbf{D}_1 \mathbf{f})_{n,m} = F_{n,m} - F_{n+1,m}$ assumes zeros boundary conditions (i.e., F is zero outside)

(d) Complete the caption of figure 2.

Solution: (c) original image ; (b) gradient across rows ; (a) gradient across columns ; (d) gradient energy

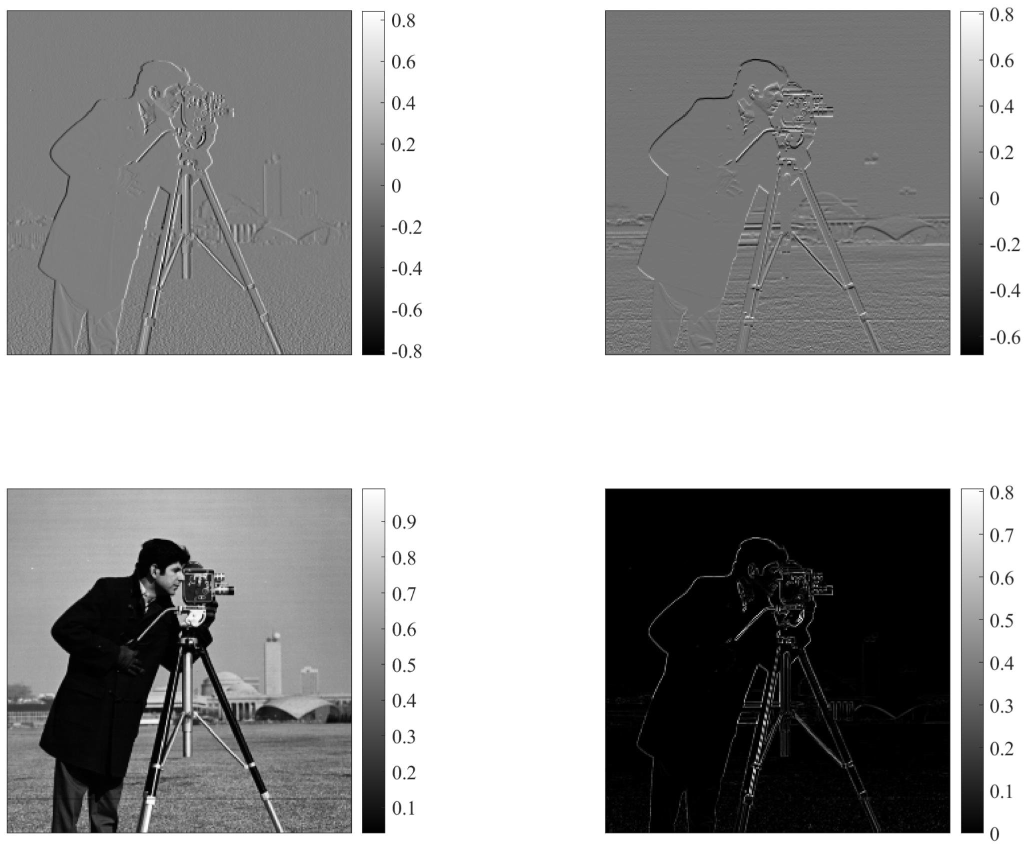

FIGURE 2 – ( ) original image ; ( ) gradient across rows ; ( ) gradient across columns ; ( ) gradient energy

5. We will consider the regularizers $\mathcal{J}_2(\mathbf{f}) = \|\nabla \mathbf{f}\|_2^2$ and $\mathcal{J}_1(\mathbf{f}) = \|\nabla \mathbf{f}\|_1$.
   (a) What kind of images will make the regularization terms $\mathcal{J}_1$ or $\mathcal{J}_2$ small ?

   Solution: Image with "small" gradients, i.e., varying slowly.

   (b) What will be the difference between minimizing the ℓ-1 and the ℓ-2 norms ?

   Solution: Jumps are more penalized with the ℓ-2 norm than with the ℓ-1 norm (quadratic vs linear penalization).

(c) What is the role of the regularization parameter α introduced in Eq. (1) ? How does the reconstructed image look like when α tends to 0 or ∞ ?

> **Solution:** It sets the compromise between the data fidelity term $\|\mathbf{s}-\mathbf{R}\mathbf{f}\|_1^2$ and the regularization term $\mathcal{J}$. Images reconstructed from small α's may suffer from the strides artefact visible in figure 1. Images reconstructed from large α's may be overly smoothed.

(d) Complete the caption of figure 3 by picking up the regularization parameter and regularization term corresponding to each of the reconstructed images.

> **Solution:** (a) ℓ-2 norm, $\alpha = 2.10^{-1}$; (c) ℓ-2 norm, $\alpha = 10^{-1}$; (e) ℓ-2 norm, $\alpha = 5.10^{-2}$; (f) ℓ-1 norm, $\alpha = 10^{-4}$; (d) ℓ-1 norm, $\alpha = 10^{-3}$; (b) ℓ-1 norm, $\alpha = 10^{-2}$.

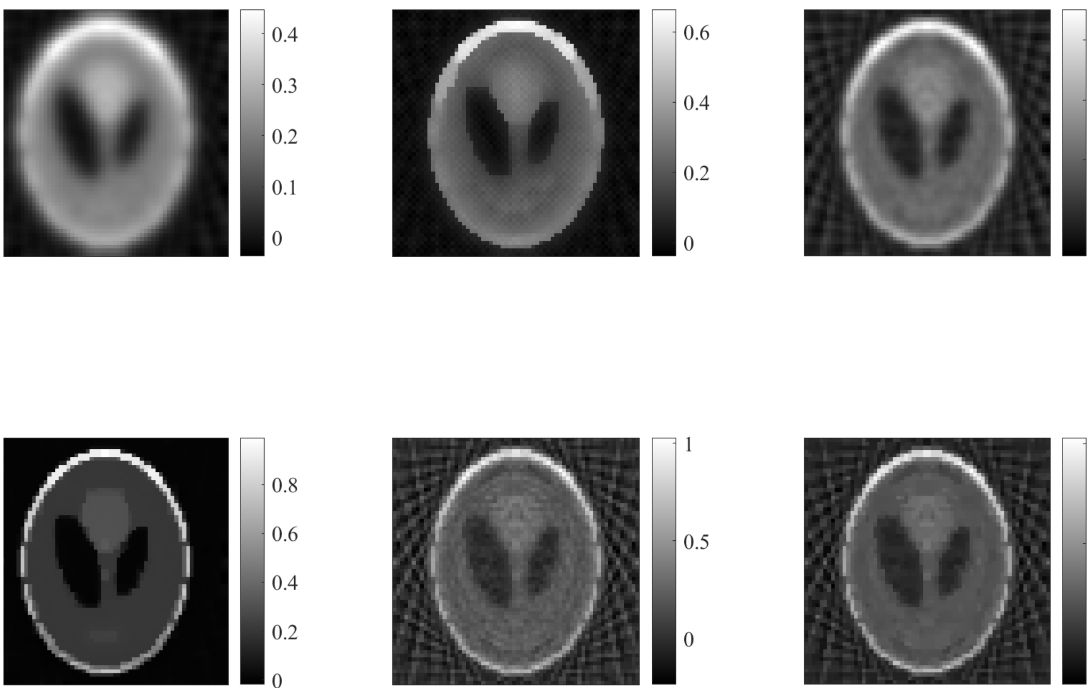

FIGURE 3 – ( ) ℓ-2 norm, $\alpha = 2.10^{-1}$; ( ) ℓ-2 norm, $\alpha = 10^{-1}$; ( ) ℓ-2 norm, $\alpha = 5.10^{-2}$; ( ) ℓ-1 norm, $\alpha = 10^{-4}$; ( ) ℓ-1 norm, $\alpha = 10^{-3}$; ( ) ℓ-1 norm, $\alpha = 10^{-2}$.

6. We finally derive the gradient of the previous two regularizers

(a) Using the result of Question 3(c), show that $\nabla \mathcal{J}_2 = 2(\mathbf{D}_1^\top \mathbf{D}_1 \mathbf{f} + \mathbf{D}_2^\top \mathbf{D}_2 \mathbf{f})$

Solution: We have

$$\nabla \mathcal{J}_2 = 2\nabla^\top \nabla \mathbf{f} = 2 \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \end{bmatrix}^\top \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \end{bmatrix} \mathbf{f} = 2[\mathbf{D}_1^\top \mathbf{D}_2^\top] \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \end{bmatrix} \mathbf{f} = 2(\mathbf{D}_1^\top \mathbf{D}_1 \mathbf{f} + \mathbf{D}_2^\top \mathbf{D}_2 \mathbf{f})$$

(b) Remarking that $|f| = \sqrt{f^2}$, show that $\nabla \|\mathbf{f}\|_1 = \mathbf{f}/\sqrt{\mathbf{f}^2}$, where division and multiplication apply component-wise.

Solution: Computing the n-th component of the gradient, we have

$$\begin{aligned} \frac{\partial}{\partial f_n} \|\mathbf{f}\|_1 &= \frac{\partial}{\partial f_n} \sum_k |f_k| = \sum_k \frac{\partial}{\partial f_n} \sqrt{f_k^2} \\ &= 2 f_n \frac{1}{2\sqrt{f_n^2}} = \frac{f_n}{\sqrt{f_n^2}} \end{aligned}$$

Therefore, $\nabla \|\mathbf{f}\|_1 = \mathbf{f}/\sqrt{\mathbf{f}^2}$.

Concluding remark As the ℓ-1 norm is not differentiable in 0, a quadratic approximation is usually considered, leading to $\nabla \|\mathbf{f}\|_1 = \mathbf{f}/\sqrt{\mathbf{f}^2 + \epsilon^2}$, where ϵ is a small constant. To get the gradient of the total variation regularization term $\mathcal{J}_1$, one can first notice that $\|\nabla \mathbf{f}\|_1 = \|\mathbf{D}_1 \mathbf{f}\|_1 + \|\mathbf{D}_2 \mathbf{f}\|_1$ and then apply the chain rule, which leads to $\mathbf{D}_1^\top (\mathbf{g}_1/\sqrt{\mathbf{g}_1^2 + \epsilon^2}) + \mathbf{D}_2^\top (\mathbf{g}_2/\sqrt{\mathbf{g}_2^2 + \epsilon^2})$, where $\mathbf{g}_1 = \mathbf{D}_1 \mathbf{f}$ and $\mathbf{g}_2 = \mathbf{D}_2 \mathbf{f}$.

Tdsi AR
21/01/2022
Durée : 15 minutes

NOM : ____________________
Prénom : ____________________

Cet énoncé est composé de 2 pages (y compris celle-ci). Merci de compléter vos nom et prénom en haut à droite de la première page et de placer vos initiales sur les pages suivantes.

Tous les documents sont autorisés.

QCM (4 points)

N.B. : **Une seule** bonne réponse est possible.

1. Quelle est la gamme de fréquences classiquement utilisée en ultrasons médicaux chez l'homme ?
   - ◯ 1-5 Hz
   - ◯ 1-5 kHz
   - ◯ 1-5 MHz
   - ◯ 1-5 GHz

2. Quelle résolution spatiale peut-on obtenir avec une sonde de 10 MHz dans les tissus biologiques ? On rappelle que la vitesse de propagation des ultrasons dans l'eau est de 1540 m/s.
   - ◯ Environ 0.1 mm
   - ◯ Exactement 1.54 mm
   - ◯ Environ 1 cm.
   - ◯ Environ 10 μm

3. Une sonde ultrasonore envoie un faisceau (durée du signal = 10 μs) pour imager un diffuseur situé à 3 cm de profondeur dans les tissus mous. Quelle est la durée **minimum** de l'acquisition pour imager le diffuseur en échographie classique ?
   - ◯ 30 μs
   - ◯ 40 μs
   - ◯ 50 μs
   - ◯ 60 μs

4. Une onde ultrasonore traverse de part en part un milieu d'atténuation 1 dB/cm/MHz. Ce milieu a pour profondeur 3 cm et l'onde ultrasonore est émise à une fréquence de 7 MHz. De combien est atténué le signal après avoir traversé le milieu ?
   - ◯ 7 dB
   - ◯ Environ 0.4 dB
   - ◯ 42 dB
   - ◯ 21 dB

5. Comment appelle t-on l'ensemble des mesures en tomography par rayons X ?
   - ◯ Un sinogramme.
   - ◯ Une radiographie.
   - ◯ Une image reconstruite.
   - ◯ Une carte d'attenuation.
   - ◯ Le bruit de Poisson.

6. On dispose d'un algorithme de reconstruction par rétroprojection filtrée. Comment réduire le temps d'acquisition ?
   - ◯ En diminuant le nombre d'angles de vue acquis.
   - ◯ En augmentant le nombre d'angles de vue acquis.
   - ◯ En diminuant le nombre de pixels du détecteur.
   - ◯ En augmentant le nombre de pixels du détecteur.
   - ◯ En utilisant un algorithme de reconstruction rapide.