Questions

1. Problem statement

   (a) What do the axes of the sinogram depicted in Fig. 1 correspond to? Which is 'view angle'? Which is 'radial position'?

   > **Solution:** horizontal axis: radial position
   > vertical axis: view angle

   (b) Let $\mathbf{Y}$ be the measured sinogram, as depicted in Fig. 1. Give the relationship between the *matrix* $\mathbf{Y}$ and the *vector* $\mathbf{y}$ defined in Eq. (1).

   > **Solution:** $\mathbf{y}$ is obtained from $\mathbf{Y}$ through vectorization.

   (c) What is the dimension of $\mathbf{Y}$?

   > **Solution:** $\mathbf{Y} \in \mathbb{R}^{L \times M}$.

   (d) We acquire measurements for $L = 20$ view angles over 180°using a detector that has $M = 89$ pixels. The unknown image is 60×60. Specify the dimensions J and I defined by Eq. (1).

   > **Solution:**
   >
   > $$I = 89 \times 20 = 1780$$
   > $$J = 60 \times 60 = 3600$$

2. We start by investigating the case $\lambda = 0$, i.e., $\mathcal{L}(\mathbf{x}) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$.

   (a) Give the gradient of the data fidelity term $g(\mathbf{x}) = \nabla\mathcal{L}(\mathbf{x})$.

   Solution: Applying standard matrix differentiation, we get

   $$g(\mathbf{x}) = -2\mathbf{A}^\top(\mathbf{y} - \mathbf{A}\mathbf{x}).$$

   (b) What is the influence of the choice for the step length τ in Eq. (??)? What if it chosen 'too small' or 'too large'?

   Solution: The algorithm will converge slowly to the solution if the step length is chosen too small, while it may diverge if it is chosen too large. Therefore, finding a 'good' step length is an important practical problem.

   (c) We look for the optimal step length that solves $\min_\tau \mathcal{L}(\mathbf{x}_k - \tau\mathbf{g}_k)$. Why is this choice optimal?

   Solution: It gives the step length such that the next iteration attains the minimum of the cost function along the search direction $\mathbf{g}_k$.

   (d) Show that minimizing $\mathcal{L}(\mathbf{x}_k - \tau\mathbf{g}_k)$ with respect to τ is equivalent to minimizing $\mathcal{L}_2(\tau) = -\tau\mathbf{g}_k^\top\mathbf{g}_k + \tau^2(\mathbf{A}\mathbf{g}_k)^\top(\mathbf{A}\mathbf{g}_k)$. Hint: expand and remove terms independent of τ. Warning: difficult but good answers will be rewarded.

   Solution: We first write

   $$\begin{aligned}\mathcal{L}(\mathbf{x}_k - \tau\mathbf{g}_k) &= \left(\mathbf{y} - \mathbf{A}(\mathbf{x}_k - \tau\mathbf{g}_k)\right)^\top\left(\mathbf{y} - \mathbf{A}(\mathbf{x}_k - \tau\mathbf{g}_k)\right)\\ &= \left(\mathbf{y} - \mathbf{A}\mathbf{x}_k\right)^\top\left(\mathbf{y} - \mathbf{A}\mathbf{x}_k - \tau\mathbf{A}\mathbf{g}_k\right)\\ &\quad - \left(\tau\mathbf{g}_k\right)^\top\left[\mathbf{A}^\top\left(\mathbf{y} - \mathbf{A}\mathbf{x}_k\right) - \tau\mathbf{A}^\top\mathbf{A}\mathbf{g}_k\right]\end{aligned}$$

   By removing the terms independent of τ, we get:

   $$\mathcal{L}_2(\tau) = -\tau\left(\mathbf{y} - \mathbf{A}\mathbf{x}_k\right)^\top\left(\mathbf{A}\mathbf{g}_k\right) - \tau\left(\mathbf{g}_k\right)^\top\left[\mathbf{A}^\top\left(\mathbf{y} - \mathbf{A}\mathbf{x}_k\right) - \tau\mathbf{A}^\top\mathbf{A}\mathbf{g}_k\right].$$

   Substituting $\mathbf{A}^\top(\mathbf{y} - \mathbf{A}\mathbf{x}_k)$ by $\mathbf{g}_k/2$, one obtain

   $$\mathcal{L}_2(\tau) = -\frac{\tau}{2}\left(\mathbf{g}_k\right)^\top\mathbf{g}_k - \frac{\tau}{2}\left(\mathbf{g}_k\right)^\top\mathbf{g}_k + \tau^2\left(\mathbf{g}_k\right)^\top\mathbf{A}^\top\mathbf{A}\mathbf{g}_k,$$

   which leads to $\mathcal{L}_2(\tau) = -\tau\mathbf{g}_k^\top\mathbf{g}_k + \tau^2(\mathbf{A}\mathbf{g}_k)^\top(\mathbf{A}\mathbf{g}_k)$ as expected.

   (e) Compute the gradient of $\mathcal{L}_2(\tau)$.

> **Solution:** $\nabla \mathcal{L}_2(\tau) = -\mathbf{g}_k^\top \mathbf{g}_k + 2\tau(\mathbf{A}\mathbf{g}_k)^\top(\mathbf{A}\mathbf{g}_k)$

(f) Show that the optimal step length is $\tau_k = (\mathbf{g}_k^\top \mathbf{g}_k)/(2\mathbf{g}_k^\top \mathbf{A}^\top \mathbf{A}\mathbf{g}_k)$. This choice leads to the so-called *steepest descent*. Hint: set the derivative of $\mathcal{L}_2$ to zero.

> **Solution:** By setting $\nabla \mathcal{L}_2(\tau)$ to zero, we obtain $2\tau(\mathbf{A}\mathbf{g}_k)^\top(\mathbf{A}\mathbf{g}_k) = \mathbf{g}_k^\top \mathbf{g}_k$ and, therefore,
>
> $$\tau_k = \frac{\mathbf{g}_k^\top \mathbf{g}_k}{2\mathbf{g}_k^\top \mathbf{A}^\top \mathbf{A}\mathbf{g}_k}.$$

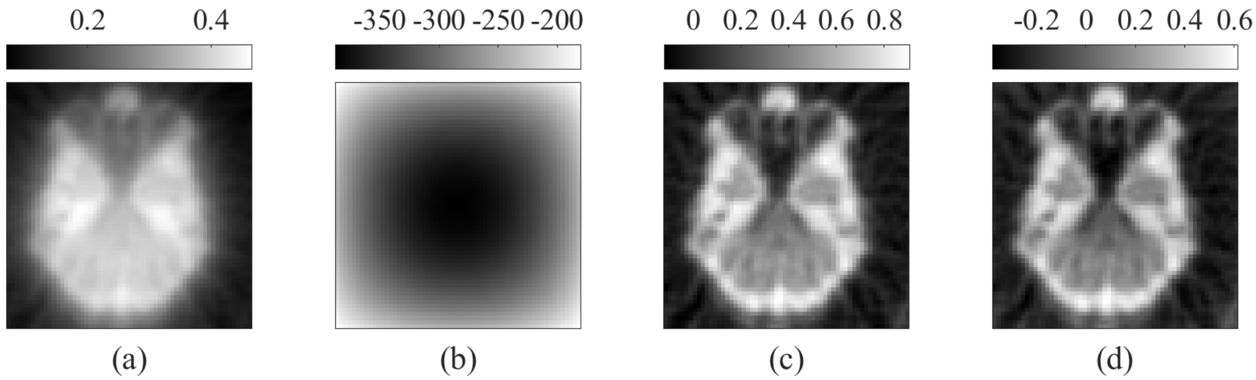


Figure 2: Reconstruction using different choices for the step length. ( ) $\tau = 1.5$; ( ) $\tau = 1$; ( ) $\tau = 0.1$; ( ) Optimal step length. In all cases, we consider 10 gradient descent iterations and set $\mathbf{x}^{(0)} = \mathbf{0}$.

(g) Complete the caption of Fig. 2 by picking up the step length corresponding to each of the reconstructed images. Justify your choices.

> **Solution:** (b) $\tau = 1.5$: the largest step as the algorithm diverges.
> (d) $\tau = 1$: fast convergence. As in (c), the image has high frequency components. However, the range of values is closer to 0 (i.e., the initial guess) than in (c).
> (a) $\tau = 0.1$: slow convergence. Compared to (c) and (d), high-frequency components are missing. It has the lowest maximum value (e.g., about 0.45) among all images.
> (c) Optimal step length: the best reconstruction with an improved range of values compared to (c).

3. We now focus on the regularizers defined by Eq. (2). We consider the case where $\mathbf{B}$ is the identity matrix and the potential function ϕ is chosen among the three candidates:

$$\phi_1(x) = x^2 \qquad \text{(quadratic)} \tag{4}$$

$$\phi_2(x) = \sqrt{x^2+\epsilon^2} - \epsilon \qquad \text{(pseudo Huber)} \tag{5}$$

$$\phi_3(x) = \frac{x^2}{x^2+\mu^2} \qquad \text{(Geman-McClure)} \tag{6}$$

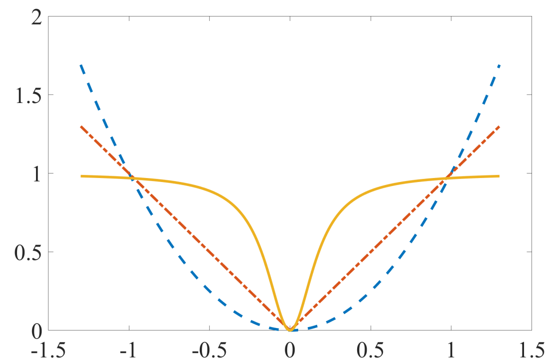


Figure 3: The three regularization potentials considered in this exercise. We set $\epsilon = 10^{-3}$ and $\mu \approx 0.18$.

(a) Annotate Fig. 3 to specify which plot corresponds to which potential.

Solution: dashed line (blue): Quadratic ϕ_1.
dotted line (red): Pseudo Huber ϕ_2.
Full line (orange): Geman-McClure ϕ_3.

(b) What kind of images will make the regularization term $\mathcal{R}$ small? Discuss the case of each of the three potential functions.

Solution:

- Quadratic ϕ_1: Pixels with large values contribute significantly more than pixels with small values. Images with few large values lead to small regularization terms.
- Pseudo Huber ϕ_2: Pixels with large values contribute more than pixels with small values, but not significantly more. Images with many small values lead to small regularization terms, but a few very large values may make it large.

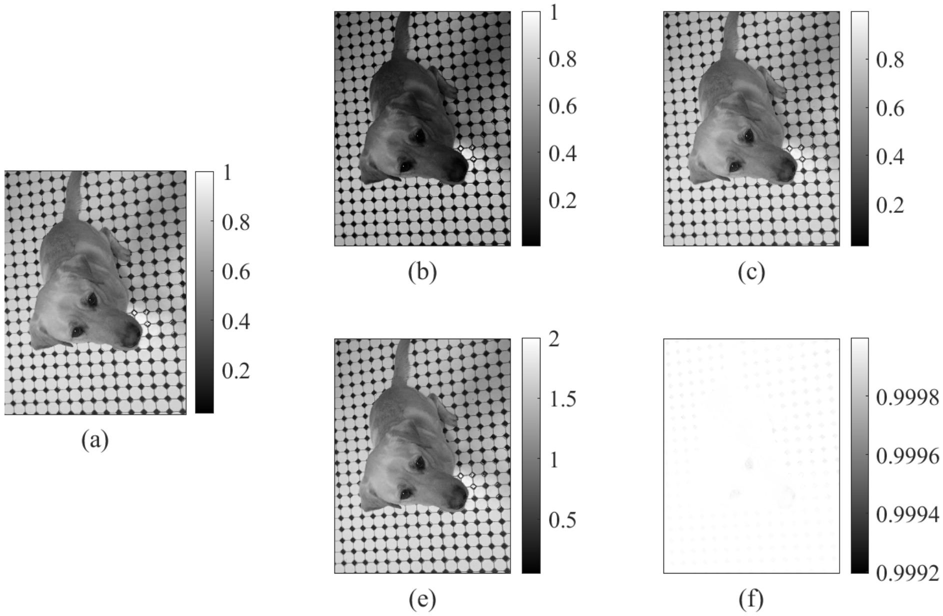


Figure 4: Regularization term for different potential functions. (a) Original image; (b–d) potential images, i.e., evaluation of the potential functions at each pixel of the original image; (e–g) potential derivative images, i.e., evaluation of the potential functions at each pixel of the original image.

- Geman-McClure ϕ_3: All pixels with large values contribute the same way. Images with many small values (e.g., many zeros) lead to small regularization terms. Images with a few pixels with very large values can also lead to small regularization terms.

(c) For each of the images (b-d) of Fig. 4, specify the potential it corresponds to? Justify.

Solution: Quadratic ϕ_1: The potential image is darker than the original image. Pseudo Huber ϕ_2: The potential image is the same as the original image. Geman-McClure ϕ_3: All pixels with values above 0.2 are set to one. The potential image is brighter than the original image.

4. We finally investigate the case $\lambda > 0$. The regularization parameter λ sets a compromise between the data fidelity term $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$ and the regularization term $\mathcal{R}$. Among solutions with a small data fidelity, it selects one with a small regularization term.

   (a) Show that the gradient of $\mathcal{L}$ is given by $\nabla\mathcal{L}(\mathbf{x}) = -2\mathbf{A}^\top(\mathbf{y} - \mathbf{A}\mathbf{x}) + \lambda\nabla\mathcal{R}(\mathbf{x})$.

   Solution: By linearity and using 2(a), $\nabla\mathcal{L}(\mathbf{x}) = -2\mathbf{A}^\top(\mathbf{y} - \mathbf{A}\mathbf{x}) + \lambda\nabla\mathcal{R}(\mathbf{x})$.

   (b) Show that the gradient of $\mathcal{R}$ only depends on the derivative of the potential in each pixel.

   Solution: By definition we compute $\frac{\partial}{\partial x_j}\mathcal{R}(\mathbf{x}) = \frac{\partial}{\partial x_j}\sum_n \phi(x_n) = \sum_n \frac{\partial}{\partial x_j}\phi(x_n) = \phi'(x_j)$. Therefore

$$\nabla\mathcal{R} = \begin{bmatrix} \phi'(x_1) \\ \vdots \\ \phi'(x_J) \end{bmatrix}$$

   (c) For each of the images (e-g) of Fig. 4, specify the potential derivative it corresponds to? Justify.

   Solution: Quadratic ϕ_1: The gradient is given by the image values.
   Pseudo Huber ϕ_2: The gradient almost one everywhere as all pixels are positive.
   Geman-McClure ϕ_3: The gradient is small where the image takes large values.

   (d) For each potential type, what pixels are most modified by a gradient descent update?

   Solution: We look at the gradient of the regularizer.
   Quadratic ϕ_1: The gradient is large where the image takes large values.
   Pseudo Huber ϕ_2: The gradient is almost one everywhere as all pixels are positive.
   Geman-McClure ϕ_3: The gradient is large where the image takes small values.

   (e) Complete the caption of Fig. 5 to indicate the potential leading to each image. Justify.

   Solution: (a) Quadratic ϕ_1; many small values in the background
   (c) Pseudo Huber ϕ_2; fewer small values in the background
   (a) Geman-McClure ϕ_3: solution with many zeros
   (d) Total variation: solution with flat regions.

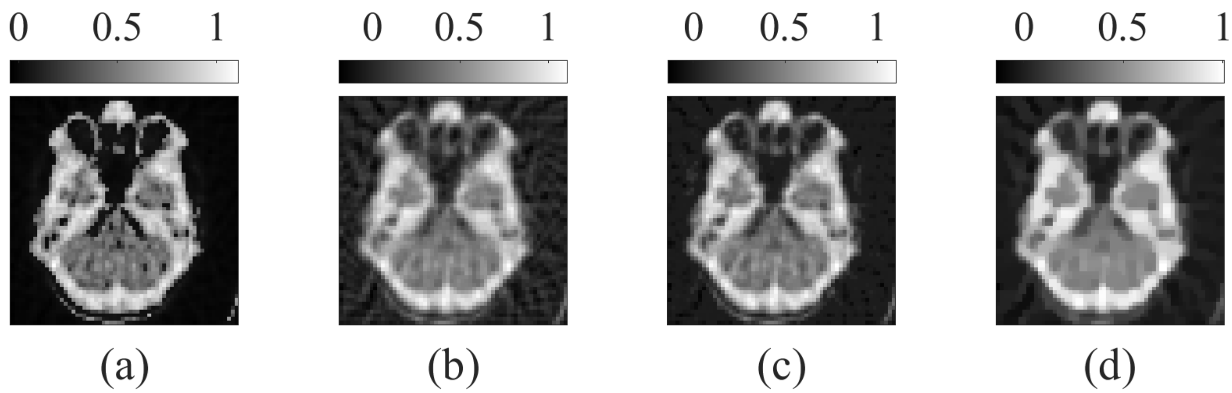


Figure 5: Reconstruction for different potential functions. ( ) Quadratic ϕ_1; ( ) Pseudo Huber ϕ_2; ( ) Geman-McClure ϕ_3; ( ) Total variation, i.e., $\phi = \phi_2$ and $\mathbf{B}$ is the gradient. We consider 20 gradient descent iterations and set $\mathbf{x}^{(0)} = \mathbf{0}$.

(f) Conclude on the choice of the regularizer term. Which is best? What are its pros and cons?

Solution: The Geman-McClure potential is better than the Pseudo Huber potential at sparcifying the solution. However, some pixel are set to zero while they take small values in the original image.

Tdsi AR
28/01/2023
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. Une onde acoustique passant d'un matériau piezo-électrique d'impédance acoustique $Z = 15$, à la peau d'impédance acoustique $Z = 1.5$, est
   - ◯ En grande partie transmise
   - ◯ En grande partie réfléchie
   - ◯ Transmise à plus de 50%
   - ◯ Réfléchie à plus de 50%

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 1 cm.
   - ◯ Environ 0.1 mm
   - ◯ Environ 10 μm
   - ◯ Exactement 1.54 mm

3. On souhaite focaliser des ondes ultrasonores au point $(x, z) = (0, 5)$ cm (cf. figure 1). Quel retard appliquer à l'élément situé en $x = 0$ pour que l'onde émise par cet élément arrive en même temps que l'onde émise par un élément situé en $x = 700$ μm ? On supposera une vitesse de propagation des ultrasons de 1540 m/s.
   - ◯ 3.2 ns
   - ◯ 32 μs
   - ◯ 65 μs
   - ◯ Aucun retard n'est nécessaire

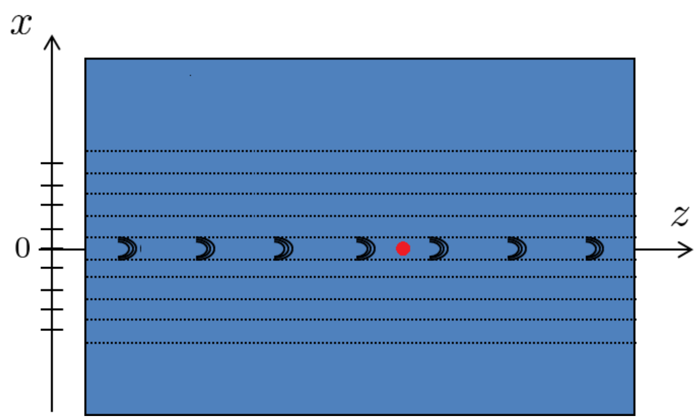


FIGURE 1 – Géométrie du problème de la question 3.

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 ?
   - ◯ 42 dB
   - ◯ 21 dB
   - ◯ 7 dB
   - ◯ On ne peut pas savoir

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

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

Tdsi AR
27th January 2022
Duration : 60 minutes

NOM : ____________________
Prénom : ____________________

This exercice has 8 pages. Fill you first name and last name on the top of the first page; put your initials on the next pages.

All documents are allowed.

Exercice (16 points)

The goal of this exercise is to solve a reconstruction problem by minimizing

$$\mathcal{L}(\mathbf{x}) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \mathcal{R}(\mathbf{x}), \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^{I\times 1}$ represents the measurement vector, $\mathbf{A} \in \mathbb{R}^{I\times J}$ the discrete Radon transform, $\mathbf{x} \in \mathbb{R}^{J\times 1}$ the attenuation (unknown) image, λ is the regularization parameter and $\mathcal{R}$ a regularization term. In particular, we will consider regularizers given by

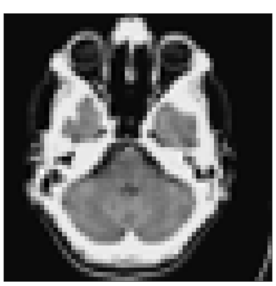
$$\mathcal{R}(\mathbf{x}) = \sum_n \phi([\mathbf{B}\mathbf{x}]_n), \tag{2}$$

where ϕ is a nonlinear function and $\mathbf{B} \in \mathbb{R}^{N\times J}$ is a linear operator (matrix).

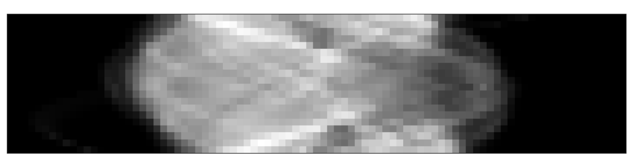
To minimize Eq. (1), we implement a gradient descent with the update rule

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \tau \mathbf{g}_k, \quad \mathbf{x}^{(0)} = \mathbf{0}, \tag{3}$$

where τ is the step length and $\mathbf{g}_k$ is the gradient of the cost function $\mathcal{L}$.



(a) unknow image



(b) measured sinogram

Figure 1: The problem is to reconstruct the unknow image $\mathbf{x}^*$ from the measured sinogram $\mathbf{A}\mathbf{x}^* + \texttt{noise}$.