Tdsi AR
24th January 2025
Duration : 50 minutes

NOM : ______________________
Prénom : ______________________

This exercice has 9 pages. Fill you first name and last name on the top of the first page; put your initials on the next pages.

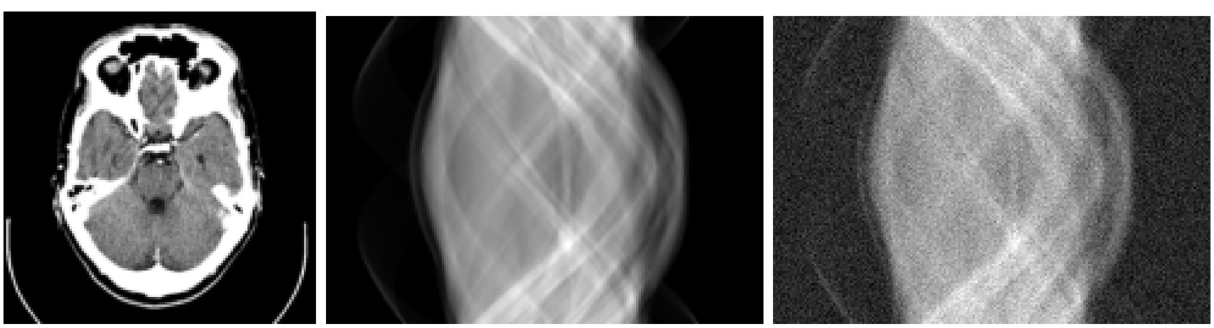
All documents are allowed.

Exercice: Maximum likelihood reconstruction (16 points)

The goal of this exercise is to recover the attenuation image $\mathbf{x} \in \mathbb{R}^{J\times 1}$ from the noisy measurements

$$\mathbf{y} = \mathcal{N}\left(\mathbf{A}\mathbf{x}\right), \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^{I\times 1}$ represents the measurement vector, $\mathbf{A} \in \mathbb{R}^{I\times J}$ is the discrete Radon transform, and $\mathcal{N}$ is the noise degradation resulting from the acquisition process. To do so, statistical methods consider that $\mathbf{y}$ is a random vector and exploit the knowledge of its probability density $p(\boldsymbol{y};\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ represents the parameters to be estimated (i.e., reconstructed).



(a) Image $\mathbf{x}$ (b) Clean sinogram $\mathbf{Ax}$ (c) Noisy sinogram $\mathbf{y}$

Figure 1: The problem is to reconstruct the unknown image $\mathbf{x}$ from the noisy sinogram $\mathcal{N}(\mathbf{Ax})$.

Maximum likelihood Maximum likelihood reconstructs the image by maximising the negative log likelihood, i.e.,

$$\mathbf{x} \in \arg\max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}) \tag{2}$$

where the negative log likelihood is given by $\ell(\boldsymbol{\theta}) = -\ln p(\mathbf{y};\boldsymbol{\theta})$. Note that this is a function of $\boldsymbol{\theta}$, given the (noisy) measurements $\mathbf{y}$.

Notations We denote the i-th entry of the vector $\mathbf{x}$ by x_i and its $\ell 2$-norm by $\|\mathbf{x}\|$. We denote the entry corresponding to the i-th row and j-th column of the matrix $\mathbf{A}$ by $A_{i,j}$. Let $\nabla\mathcal{L}$ be the gradient of a function $\mathcal{L} : \mathbb{R}^I \to \mathbb{R}$

$$\nabla\mathcal{L} = \begin{bmatrix} \frac{\partial}{\partial x_1}\mathcal{L} \\ \vdots \\ \frac{\partial}{\partial x_I}\mathcal{L} \end{bmatrix}$$

Distributions A random variable Y that follows a Gaussian distribution $\mathcal{G}(\mu, \sigma^2)$ has probability density

$$p(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right], \quad y \in \mathbb{R}, \tag{3}$$

where $\mu \in \mathbb{R}$ is the expectation and $\sigma^2 \in \mathbb{R}_+$ is the variance.

A random variable Y that follows a Poisson distribution $\mathcal{P}(\mu)$ has probability

$$p(y; \mu) = \frac{\mu^y \exp(-\mu)}{y!} \tag{4}$$

to take the value $y \in \mathbb{N}$, where $\mu \in \mathbb{N}$ is the expectation (and also the variance).

Questions

1. Problem statement

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

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

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

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

(c) We acquire measurements for 100 view angles over 180°using a detector that has 182 pixels. What is the dimension of $\mathbf{Y}$?

> **Solution:** $\mathbf{Y} \in \mathbb{R}^{100\times 182}$.

(d) The unknown image is 128×128. Specify the dimensions I and J defined by Eq. (1).

> **Solution:**
>
> $$I = 100 \times 182 = 18200$$
> $$J = 128 \times 128 = 16384$$

2. We start by investigating the case where each of the component of the measurement vector are independent identically distributed (i.i.d.) Gaussian variables, i.e.,

$$y_i = \mathcal{G}\left(\mu_i = \mathbf{a}_i^\top \mathbf{x}, \sigma^2\right), \quad 1 \leq i \leq I, \tag{5}$$

where $\mathbf{a}_i \in \mathbb{R}^J$ represents the i-th row of $\mathbf{A}$ and σ^2 characterises the acquisition device.

(a) Give the expression of the likelihood of the image $\mathbf{x}$ given only the i-th component of the measurement vector is available. Note: We assume that $\mathbf{a}_i$ and σ are known.

Solution: From the definition

$$\ell(\mathbf{x}) = -\ln p(y_i; \mathbf{x}) = -\ln\left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_i - \mathbf{a}_i^\top \mathbf{x})^2}{2\sigma^2}\right]\right), \tag{6}$$

$$= \ln C + \frac{(y_i - \mathbf{a}_i^\top \mathbf{x})^2}{2\sigma^2} \tag{7}$$

where $C = \sqrt{2\pi\sigma^2}$ is a constant independent of $\mathbf{x}$.

(b) Assuming i.i.d variables, what is the joint probability $p(\mathbf{y}; \boldsymbol{\theta}) = p(y_1, \ldots, y_I; \boldsymbol{\theta})$ that the measurements take value $\mathbf{y}$?

Solution:

$$p(y_1, \ldots, y_I; \boldsymbol{\theta}) = \prod_i p(y_i; \boldsymbol{\theta}) = C^I \exp\left(-\sum_i \frac{(y_i - \mathbf{a}_i^\top \mathbf{x})^2}{\sqrt{2\pi\sigma^2}}\right) \tag{8}$$

(c) Show that the negative log-likelihood of the image $\mathbf{x}$ given the full measurement vector $\mathbf{y}$ is given by $\ell(\mathbf{x}) = I \ln C + \frac{1}{2\sigma^2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2$, with $C = \sqrt{2\pi\sigma^2}$.

Solution:

$$\ell(\mathbf{x}) = -\ln p(\mathbf{y}; \mathbf{x}) = I \ln C + \frac{1}{2\sigma^2} \sum_i (y_i - \mathbf{a}_i^\top \mathbf{x})^2 \tag{9}$$

$$= I \ln C + \frac{1}{2\sigma^2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 \tag{10}$$

(d) By differentiation of the likelihood, derive the expression of the maximum likelihood solution (i.e., the solution of Eq. (2)).

Solution:

$$\nabla \ell(\mathbf{x}) = \nabla \frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 = -\frac{1}{2\sigma^2} \mathbf{A}^\top (\mathbf{y} - \mathbf{A}\mathbf{x}) \tag{11}$$

By setting the derivative to zero, we have

$$\mathbf{0} = \mathbf{A}^\top (\mathbf{y} - \mathbf{A}\mathbf{x}) \Leftrightarrow \mathbf{A}^\top \mathbf{A}\mathbf{x} = \mathbf{A}^\top \mathbf{y} \tag{12}$$

Assuming $\mathbf{A}^\top \mathbf{A}$ is invertible, the solution is given by

$$\mathbf{x} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}\mathbf{y} \tag{13}$$

(e) How is this solution also referred to as?

Solution: It is often referred to as the pseudoinverse solution.

3. We now consider the case where each of the component of the measurement vector are i.i.d. Poisson variables, i.e.,

$$y_i = \mathcal{P}\left(\mu_i = \mathbf{a}_i^\top \mathbf{x}\right), \quad 1 \le i \le I, \tag{14}$$

   (a) Give the expression of the likelihood of the image $\mathbf{x}$ given only the i-th component of the measurement vector is available. Note: We assume that $\mathbf{a}_i$ is known.

Solution: From the definition

$$\begin{aligned}
\ell(\mathbf{x}) = -\ln p(y_i;\mathbf{x}) &= -\ln\left(\frac{(\mathbf{a}_i^\top \mathbf{x})^{y_i} \exp\left(-\mathbf{a}_i^\top \mathbf{x}\right)}{y_i!}\right), \\
&= C + \mathbf{a}_i^\top \mathbf{x} - \ln(\mathbf{a}_i^\top \mathbf{x})^{y_i} \\
&= C + \mathbf{a}_i^\top \mathbf{x} - y_i \ln \mathbf{a}_i^\top \mathbf{x}
\end{aligned}$$

where $C = \ln y_i!$ is a constant independent of $\mathbf{x}$.

   (b) Show that the negative log-likelihood of the image $\mathbf{x}$ given the full measurement vector $\mathbf{y}$ is given by $\ell(\mathbf{x}) = IC + \sum_i \mathbf{a}_i^\top \mathbf{x} - y_i \ln(\mathbf{a}_i^\top \mathbf{x})$, with $C = \ln y_i!$.

Solution:

$$\begin{aligned}
\ell(\mathbf{x}) = -\ln p(\mathbf{y};\mathbf{x}) &= -\sum_i \ln p(y_i;\mathbf{x}) \\
&= \sum_i C + \mathbf{a}_i^\top \mathbf{x} - y_i \ln \mathbf{a}_i^\top \mathbf{x} \\
&= IC + \sum_i \mathbf{a}_i^\top \mathbf{x} - \ln(\mathbf{a}_i^\top \mathbf{x})^{y_i} \\
&= IC + \sum_i \mathbf{a}_i^\top \mathbf{x} - y_i \ln(\mathbf{a}_i^\top \mathbf{x})
\end{aligned}$$

   (c) Show that the derivative of the negative log-likelihood can be written $\nabla\ell(\mathbf{x}) = \sum_i \nabla(\mathbf{a}_i^\top \mathbf{x}) - \nabla(y_i \ln(\mathbf{a}_i^\top \mathbf{x}))$.

Solution:

$$\begin{aligned}
\nabla\ell(\mathbf{x}) &= \nabla\left[IC + \sum_i \mathbf{a}_i^\top \mathbf{x} - y_i \ln(\mathbf{a}_i^\top \mathbf{x})\right] \\
&= \sum_i \nabla(\mathbf{a}_i^\top \mathbf{x}) - \nabla(y_i \ln(\mathbf{a}_i^\top \mathbf{x}))
\end{aligned}$$

   (d) Show that $\nabla(\mathbf{a}_i^\top \mathbf{x}) = \mathbf{a}_i$

Solution:

$$\frac{\partial}{\partial x_j}(\mathbf{a}_i^\top \mathbf{x}) = \frac{\partial}{\partial x_j}\sum_k (a_{i,k}x_k) = a_{i,j}$$

(e) Show that $\nabla(y_i \ln(\mathbf{a}_i^\top \mathbf{x})) = y_i \frac{1}{\mathbf{a}_i^\top \mathbf{x}}\mathbf{a}_i$

Solution:

$$\begin{aligned}\frac{\partial}{\partial x_j} y_i \ln(\mathbf{a}_i^\top \mathbf{x}) &= y_i \frac{1}{\mathbf{a}_i^\top \mathbf{x}}\frac{\partial}{\partial x_j}\left(\mathbf{a}_i^\top \mathbf{x}\right) \\ &= y_i \frac{1}{\mathbf{a}_i^\top \mathbf{x}} a_{i,j}\end{aligned}$$

(f) Show that the derivative of the negative log-likelihood can be written $\mathbf{A}^\top \mathbf{1} - \mathbf{A}^\top \frac{\mathbf{y}}{\mathbf{A}\mathbf{x}}$, where vector division applies component-wise.

Solution:

$$\begin{aligned}\nabla \ell(\mathbf{x}) &= \sum_i \nabla(\mathbf{a}_i^\top \mathbf{x}) - \nabla(y_i \ln(\mathbf{a}_i^\top \mathbf{x})) = \sum_i \mathbf{a}_i - \mathbf{a}_i \frac{y_i}{\mathbf{a}_i^\top \mathbf{x}} \\ &= \mathbf{A}^\top \mathbf{1} - \sum_i \mathbf{A}^\top \frac{y_i}{\mathbf{a}_i^\top \mathbf{x}} = \mathbf{A}^\top \mathbf{1} - \mathbf{A}^\top \frac{\mathbf{y}}{\mathbf{A}\mathbf{x}}\end{aligned}$$

(g) How the maximum likelihood solution can be obtained from the gradient derived in the previous question? Discuss the pros and cons.

Solution: By gradient descent. This is a simple approach but the gradient is not defined in $\mathbf{0}$.

4. To find the points of a function $\mathcal{F}$ that satisfy $\mathcal{F}(\mathbf{x}) = \mathbf{x}$, one can consider the fixed-point iteration $\mathbf{x}^{(k+1)} = \mathcal{F}(\mathbf{x}^{(k)})$. Such solutions are called fixed-points.

   (a) In the case of Gaussian noise, show that $\mathbf{x} = \frac{\mathbf{A}^\top \mathbf{y}}{\mathbf{A}^\top \mathbf{A}\mathbf{x}} \odot \mathbf{x}$, where $\odot$ represents the component-wise product of two vectors.

   Solution: First order optimality imposes $\nabla \ell(\mathbf{x}) = \mathbf{0}$, leading to

$$\begin{aligned}(\mathbf{A}^\top \mathbf{A}\mathbf{x} - \mathbf{A}^\top \mathbf{y}) \odot \mathbf{x} &= \mathbf{0} \\ \mathbf{A}^\top \mathbf{A}\mathbf{x} \odot \mathbf{x} &= \mathbf{A}^\top \mathbf{y} \odot \mathbf{x} \\ \mathbf{x} &= \frac{\mathbf{A}^\top \mathbf{y}}{\mathbf{A}^\top \mathbf{A}\mathbf{x}} \odot \mathbf{x}\end{aligned}$$

   (b) Derive an iteration that is hoped to converge to ML solution in the case of Gaussian noise. Hint: This is a multiplicative update.

   Solution:

$$\mathbf{x}^{(k+1)} = \frac{\mathbf{A}^\top \mathbf{y}}{\mathbf{A}^\top \mathbf{A}\mathbf{x}^{(k)}} \odot \mathbf{x}^{(k)}$$

   (c) In the case of Poisson noise, show that $\mathbf{x} = \frac{1}{\mathbf{A}^\top \mathbf{1}} \odot \mathbf{A}^\top \frac{\mathbf{y}}{\mathbf{A}\mathbf{x}} \odot \mathbf{x}$

   Solution: First order optimality imposes $\nabla \ell(\mathbf{x}) = \mathbf{0}$, leading to

$$\begin{aligned}(\mathbf{A}^\top \mathbf{1} - \mathbf{A}^\top \frac{\mathbf{y}}{\mathbf{A}\mathbf{x}}) \odot \mathbf{x} &= \mathbf{0} \\ \mathbf{A}^\top \mathbf{1} \odot \mathbf{x} &= \mathbf{A}^\top \frac{\mathbf{y}}{\mathbf{A}\mathbf{x}} \odot \mathbf{x} \\ \mathbf{x} &= \frac{1}{\mathbf{A}^\top \mathbf{1}} \odot \mathbf{A}^\top \frac{\mathbf{y}}{\mathbf{A}\mathbf{x}} \odot \mathbf{x}\end{aligned}$$

   (d) Derive an iteration that is hoped to converge to ML solution in the case of Poisson noise. Hint: This is a multiplicative update.

   Solution:

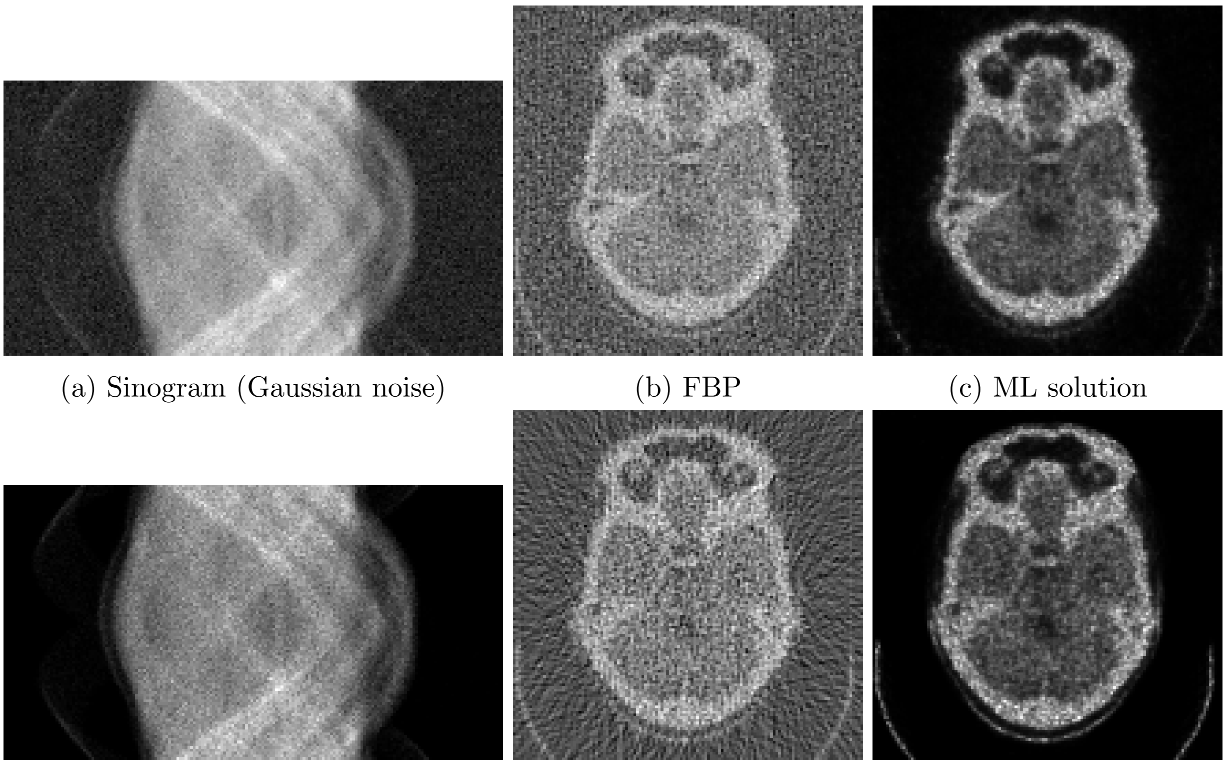
$$\mathbf{x}^{(k+1)} = \frac{\mathbf{x}^{(k)}}{\mathbf{A}^\top \mathbf{1}} \odot \mathbf{A}^\top \frac{\mathbf{y}}{\mathbf{A}\mathbf{x}^{(k)}}$$

Concluding remark Provided that the entries of $\mathbf{A}$ are positive, the multiplicative updates derived in this section guarantee that the solution remains positive (see Figure 2).

To enforce the positivity of the solution, i.e., to solve

$$\mathbf{x} \in \underset{\mathbf{x}>\mathbf{0}}{\arg\max}\, \ell(\mathbf{x}), \tag{15}$$

one can also consider the parameterisation $\mathbf{x} = \exp(\mathbf{z})$, where exponentiation applies component-wise. Using the chain rule, the gradient of the likelihood with respect to $\mathbf{z}$ is given by $\nabla_{\mathbf{x}}\ell(\mathbf{x}) \odot \mathbf{x}$.



(a) Sinogram (Gaussian noise) (b) FBP (c) ML solution

(d) Sinogram (Poisson noise) (e) FBP (f) ML solution

Figure 2: Reconstruction under Gaussian and Poisson noise using filtered back projection (FBP) and maximum likelihood (ML). The code used to compute the solution is available at https://github.com/nducros/image-reconstruction.

Tdsi AR
24 janvier 2025
Durée : 10 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.

Seules vos notes de cours sont autorisées.

QCM (4 points)

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

1. Le signal mesuré par un détecteur de rayons X est proportionnel
   - ◯ à la texture et la forme de l'objet
   - ◯ à l'atténuation de l'objet
   - ◯ au nombre de photons incidents sur l'objet
   - ◯ à l'angle de vue optimal

2. En rayons X, la correction "flat field" permet de
   - ◯ compenser les défauts du détecteur
   - ◯ réduire le bruit aléatoire du détecteur
   - ◯ s'affranchir des rayons X réfléchis sur des parois extérieures
   - ◯ s'affranchir du courant d'obscurité du détecteur

3. Comment réduire le temps d'une acquisition tomographique ?
   - ◯ En augmentant le nombre d'angles de vue.
   - ◯ En diminuant le nombre de pixels du détecteur.
   - ◯ En augmentant le nombre de pixels du détecteur.
   - ◯ En diminuant le nombre d'angles de vue.
   - ◯ En utilisant un algorithme de reconstruction rapide.

4. Quelle résolution spatiale peut-on atteindre avec une sonde de 6 MHz dans les tissus biologiques ? Rappel : la vitesse de propagation des ultrasons dans l'eau est de 1540 m/s.
   - ◯ Précisément 6 µm.
   - ◯ Environ 250 µm.
   - ◯ Environ 3.9 mm.

- ◯ On ne peut pas savoir.

5. La tomographie par rayons X consiste à reconstruire l'image de l' . . . . . . d'un objet à partir d'un ensemble de . . . . . . acquises pour . . . . . .
   - ◯ amplitude / prédictions / pour le même angle de vue.
   - ◯ amplitude / projections / pour 180 angles de vue.
   - ◯ atténuation / prédictions / pour 360 angles de vue.
   - ◯ atténuation / projections / différents angles de vue bien choisis.
   - ◯ atténuation / prédictions / pour 180 angles de vue.
   - ◯ amplitude / projections / différents angles de vue bien choisis.

6. En tomographie par ultrasons, le gel est utilisé pour
   - ◯ ne pas abimer la sonde
   - ◯ protéger la peau des effets délétères des ultrasons
   - ◯ pour limiter les ruptures d'impédance acoustique
   - ◯ nettoyer les impuretés à la surface de la peau