

Méthodes variationnelles pour l'imagerie

Sandrine ANTHOINE, Caroline CHAUX

I2M - UMR CNRS 7373, Aix-Marseille Université, France

GEOMDATA, Fréjus, 10-14 Sep. 2018



Plan de la séance

INTRODUCTION ET EXEMPLES

Introduction

Exemples

L'INVERSION, UN PROBLÈME DIFFICILE

Problèmes mal posés

Résolution par (pseudo-)inversion

SVD tronquée

Estimateur de Wiener

RÉGULARISATION

Régularisation de Tikhonov

La régularisation d'un point de vue variationnel et statistique

MAP

CONCLUSIONS

Problèmes Inverses :

Résoudre un problème inverse c'est

- ▶ estimer un objet y
- ▶ à partir de mesures indirectes z et
- ▶ d'information sur l'opérateur "direct" A qui transforme l'objet en les mesures $z = A(y)$.

On peut étudier les problèmes inverses selon plusieurs points de vue : EDP, statistique...

→ Ici, on étudie les problèmes inverses en traitement du signal et des images.

Problèmes Inverses :

Résoudre un problème inverse c'est

- ▶ estimer un objet y
- ▶ à partir de mesures indirectes z et
- ▶ d'information sur l'opérateur "direct" A qui transforme l'objet en les mesures $z = A(y)$.

On peut étudier les problèmes inverses selon plusieurs points de vue : EDP, statistique...

→ Ici, on étudie les problèmes inverses en traitement du signal et des images.

Modèle :

$$z = \mathcal{D}_\alpha(Ly)$$

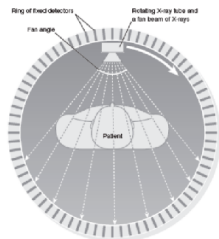
- ▶ z : mesures/observations
- ▶ y : signal original
- ▶ L : opérateur linéaire
- ▶ \mathcal{D}_α : perturbation de paramètre α .

Objectif :

Estimer y à partir des mesures z et d'informations sur \mathcal{D}_α et L .

Exemples

Projection tomographique

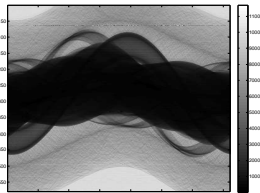
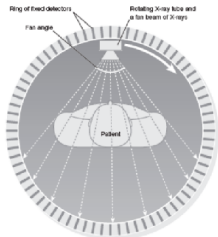


Images : Y. Boursier

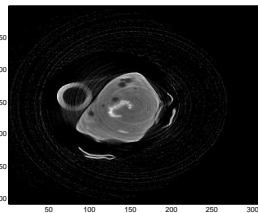
► $z = D_\alpha(Ly),$

Exemples

Projection tomographique



$$z = \mathcal{P}_\alpha(Ly)$$



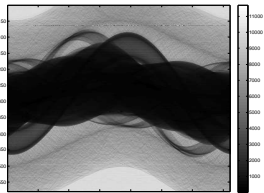
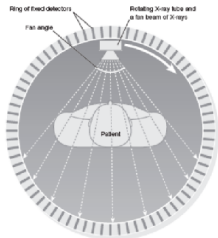
y

Images : Y. Boursier

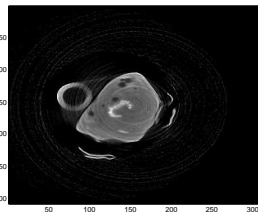
- ▶ $z = D_\alpha(Ly),$
- ▶ L l'opérateur de projection tomographique

Exemples

Projection tomographique



$$z = \mathcal{P}_\alpha(Ly)$$



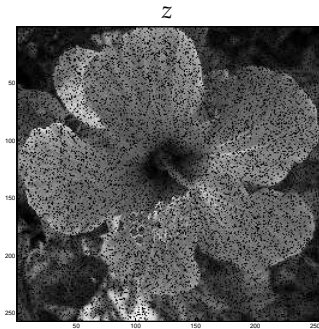
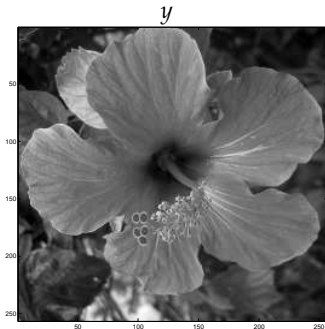
y

Images : Y. Boursier

- ▶ $z = D_\alpha(Ly)$,
- ▶ L l'opérateur de projection tomographique
- ▶ $D_\alpha = \mathcal{P}_\alpha(\cdot)$: bruit de Poisson.

Exemples

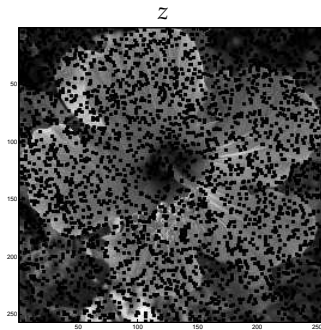
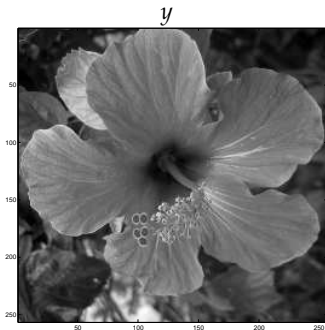
Inpainting



- ▶ $z = D_\alpha(Ly)$,
- ▶ L est un masque : $Ly(k_1, k_2) = y(k_1, k_2)$ où 0.
- ▶ $D_\alpha = \text{Id}$: pas de bruit (ici).

Exemples

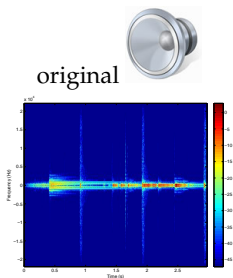
Inpainting



- ▶ $z = D_\alpha(Ly),$
- ▶ L est un masque : $Ly(k_1, k_2) = y(k_1, k_2)$ où 0.
- ▶ $D_\alpha = \text{Id}$: pas de bruit (ici).

Séparation de sources

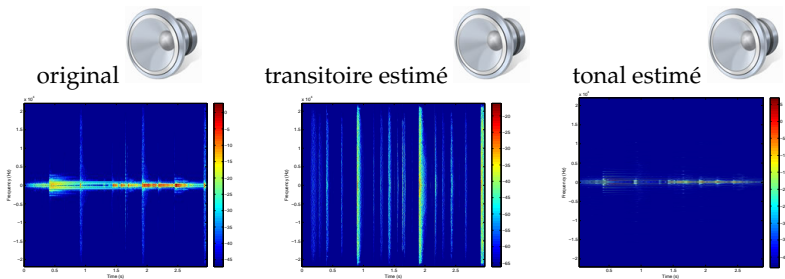
Extraction tonal versus transitoire.



P. Balasz, M. Dörfler, M. Kowalski, B. Torrèsani

Séparation de sources

Extraction tonal versus transitoire.

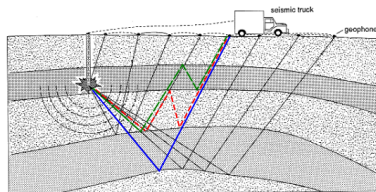


P. Balasz, M. Dörfler, M. Kowalski, B. Torrèsani

Données sismiques

Données sismiques

Acquisition

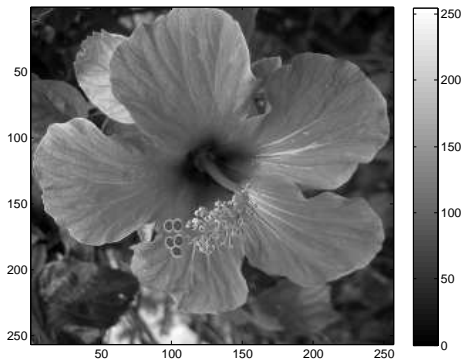


Principes de la propagation des ondes sismiques, avec des réflexions sur les différentes couches de terrain et l'acquisition des données. Trait plein **primaire**; pointillés **rouges** et **verts** : nuisances dues aux réflexions multiples.

- ▶ Le **primaire** est le signal qu'on cherche à estimer.
- ▶ Les **multiples** sont les signaux qu'on souhaite éliminer.

Notations

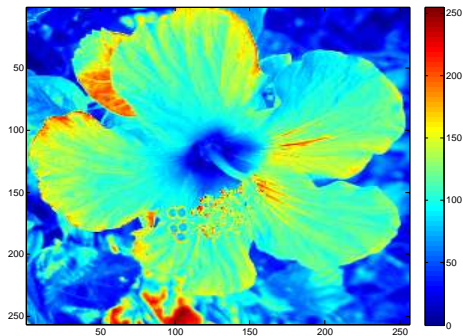
Matrices et images



Une image en niveau de gris $M \in \mathbb{R}^{N_1 \times N_2}$

Notations

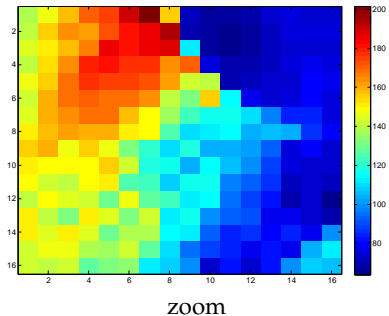
Matrices et images



Même matrice $M \in \mathbb{R}^{N_1 \times N_2}$ affichée avec une palette de couleur

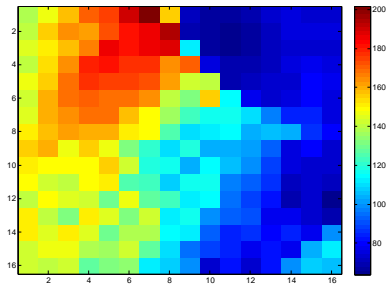
Notations

Matrices et images



Notations

Matrices et images



zoom



image vectorisée $M \in \mathbb{R}^N$ avec $N = N_1 N_2$

Exemples

Débruitage : bruit additif.



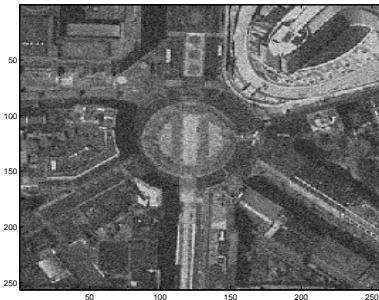
Image originale

Exemples

Débruitage : bruit additif.



Image originale



$\sigma = 20$

- ▶ $z = D_\alpha(y) = y + n$
- ▶ $L = Id$
- ▶ $(n[k_1, k_2])_{k_1, k_2}$ sont des variables aléatoires i.i.d Gaussiennes.

Exemples

Débruitage : bruit de Poisson.

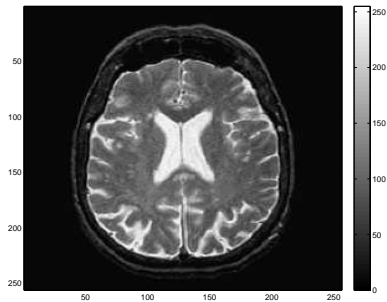


Image originale

Exemples

Débruitage : bruit de Poisson.

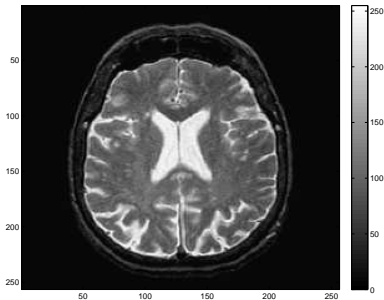
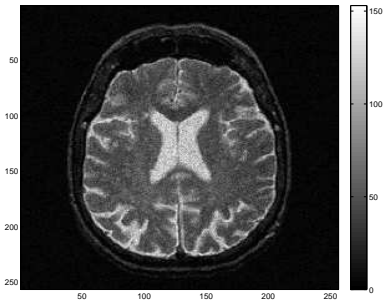


Image originale



$\alpha = 0.5$

- ▶ $z = D_\alpha(y) = \mathcal{P}_\alpha(y)$,
- ▶ $L = Id$
- ▶ $\mathcal{P}_\alpha(y) : [k_1, k_2] \mapsto \mathcal{P}(\alpha y[k_1, k_2])$ est une variable aléatoire suivant une loi de Poisson de paramètre $\alpha y[k_1, k_2]$.

Exemples

Débruitage : bruit de Poisson.

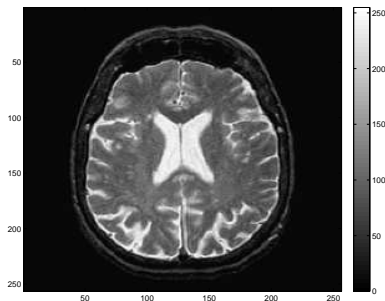
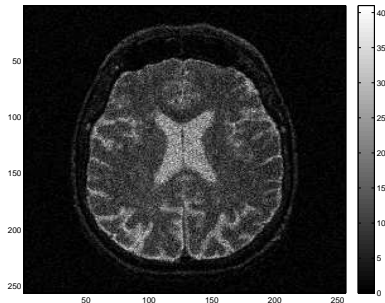


Image originale

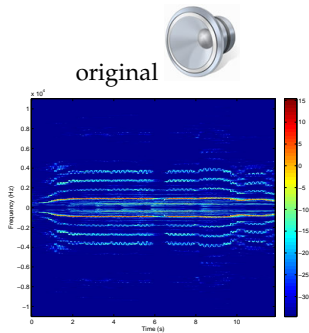


$\alpha = 0.1$

- ▶ $z = D_\alpha(y) = \mathcal{P}_\alpha(y)$,
- ▶ $L = Id$
- ▶ $\mathcal{P}_\alpha(y) : [k_1, k_2] \mapsto \mathcal{P}(\alpha y[k_1, k_2])$ est une variable aléatoire suivant une loi de Poisson de paramètre $\alpha y[k_1, k_2]$.

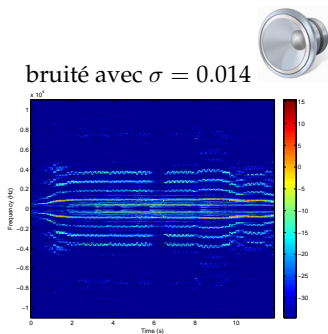
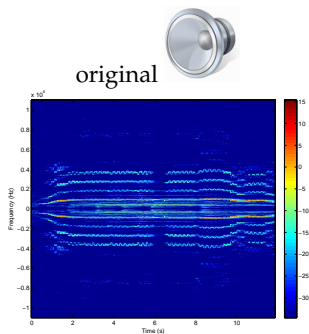
Exemples

Débruitage : bruit additif.



Exemples

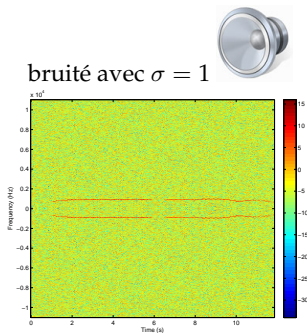
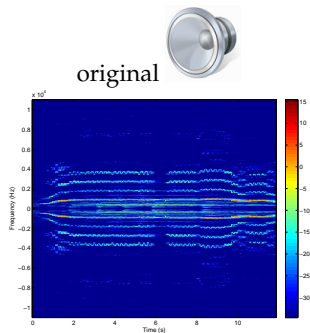
Débruitage : bruit additif.



- ▶ $z = D_\alpha(y) = y + n$
- ▶ $L = Id$
- ▶ $(n[k])_k$ sont des variables aléatoires i.i.d Gaussiennes.

Exemples

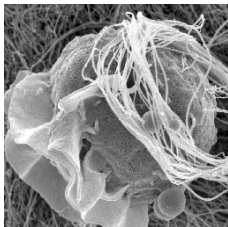
Débruitage : bruit additif.



- ▶ $z = D_\alpha(y) = y + n$
- ▶ $L = Id$
- ▶ $(n[k])_k$ sont des variables aléatoires i.i.d Gaussiennes.

Exemples

Déconvolution en microscopie



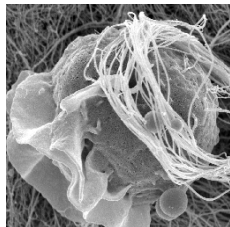
y

► $z = D_\alpha(Ly),$

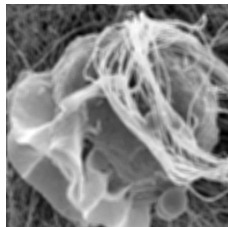
Images : C. Chaux

Exemples

Déconvolution en microscopie



y



$h \star y$

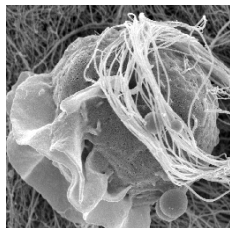
Images : C. Chaux

- ▶ $z = D_\alpha(Ly)$,
- ▶ $Ly = h \star y$: produit de convolution

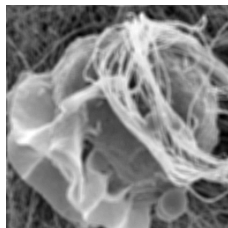
$$(h \star y)(s) = \sum_k h_k y_{s-k}$$

Exemples

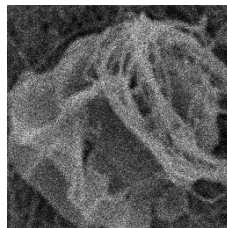
Déconvolution en microscopie



y



$h \star y$



$z = \mathcal{P}_\alpha(h \star y)$

Images : C. Chaux

- ▶ $z = D_\alpha(Ly)$,
- ▶ $Ly = h \star y$: produit de convolution

$$(h \star y)(s) = \sum_k h_k y_{s-k}$$

- ▶ $D_\alpha = \mathcal{P}_\alpha(\cdot)$: bruit de Poisson.

Plan de la séance

INTRODUCTION ET EXEMPLES

Introduction

Exemples

L'INVERSION, UN PROBLÈME DIFFICILE

Problèmes mal posés

Résolution par (pseudo-)inversion

SVD tronquée

Estimateur de Wiener

RÉGULARISATION

Régularisation de Tikhonov

La régularisation d'un point de vue variationnel et statistique

MAP

CONCLUSIONS

Problèmes mal posés

La résolution d'un problème (non-)linéaire de la forme

$$z = Ly$$

est dite **bien posée** si la solution vérifie les

Problèmes mal posés

La résolution d'un problème (non-)linéaire de la forme

$$z = Ly$$

est dite **bien posée** si la solution vérifie les

Conditions de *Hadamard*

1. Il existe au moins une solution

$$\Leftrightarrow L \in \mathbb{R}^{M \times N} \text{ linéaire : } L(\mathbb{R}^N) = \mathbb{R}^M.$$

2. La solution est unique

$$\Leftrightarrow L \text{ linéaire : } \ker L = \{0\}.$$

3. La solution est stable :

$$\forall (z, z') \in (\mathbb{R}^M)^2,$$

$$\lim_{\|z-z'\| \rightarrow 0} \|\hat{y}(z) - \hat{y}(z')\| = 0.$$

Problèmes mal posés

La résolution d'un problème (non-)linéaire de la forme

$$z = Ly$$

est dite **bien posée** si la solution vérifie les

Conditions de *Hadamard*

1. Il existe au moins une solution

$$\Leftrightarrow L \in \mathbb{R}^{M \times N} \text{ linéaire : } L(\mathbb{R}^N) = \mathbb{R}^M.$$

2. La solution est unique

$$\Leftrightarrow L \text{ linéaire : } \ker L = \{0\}.$$

3. La solution est stable :

$$\forall (z, z') \in (\mathbb{R}^M)^2,$$

$$\lim_{\|z-z'\| \rightarrow 0} \|\hat{y}(z) - \hat{y}(z')\| = 0.$$

Tout problème pour lequel au moins une de ces conditions n'est pas réalisée est dit **mal-posé**.

Déconvolution par inversion

Modèle direct :

$$z = h \star y + n$$

avec $(n[k])_k$ variables aléatoires i.i.d Gaussiennes.

Solution par inversion :

Déconvolution par inversion

Modèle direct :

$$z = h \star y + n$$

avec $(n[k])_k$ variables aléatoires i.i.d Gaussiennes.

Solution par inversion :

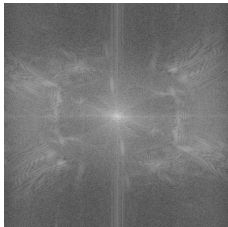
Rappel : $\hat{h}[k_1, k_2] = \sum_{\xi_1, \xi_2} h[\xi_1, \xi_2] e^{-\frac{i2\pi k_1 \xi_1}{N_1}} e^{-\frac{i2\pi k_2 \xi_2}{N_2}}$.

La transformée de Fourier d'une image

Image originale y



$|\hat{y}|$



La transformée de Fourier d'une image

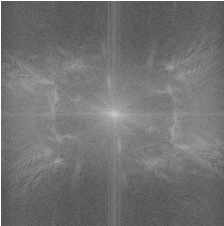
Image originale y



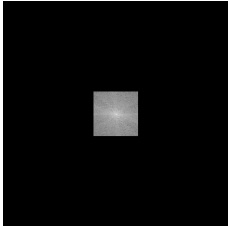
Reconstruction



$|\hat{y}|$



Basses fréquences



La transformée de Fourier d'une image

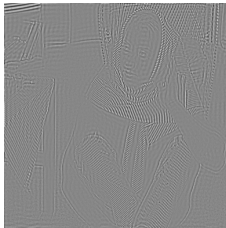
Image originale y



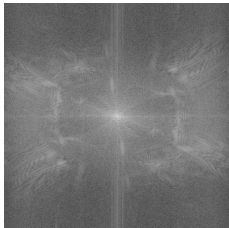
Reconstruction



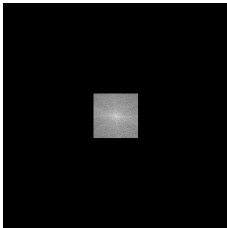
Reconstruction



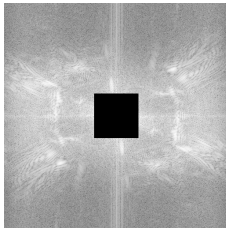
$|\hat{y}|$



Basses fréquences



Hautes fréquences



Déconvolution par inversion

Modèle direct :

$$z = h \star y + n$$

avec $(n[k])_k$ variables aléatoires i.i.d Gaussiennes.

Solution par inversion :

$$\text{Rappel : } \hat{h}[k_1, k_2] = \sum_{\xi_1, \xi_2} h[\xi_1, \xi_2] e^{-\frac{i2\pi k_1 \xi_1}{N_1}} e^{-\frac{i2\pi k_2 \xi_2}{N_2}} .$$

$$z = h \star y$$

Déconvolution par inversion

Modèle direct :

$$z = h \star y + n$$

avec $(n[k])_k$ variables aléatoires i.i.d Gaussiennes.

Solution par inversion :

$$\text{Rappel : } \widehat{h}[k_1, k_2] = \sum_{\xi_1, \xi_2} h[\xi_1, \xi_2] e^{-\frac{i2\pi k_1 \xi_1}{N_1}} e^{-\frac{i2\pi k_2 \xi_2}{N_2}} .$$

$$z = h \star y$$

$$\widehat{z} = \widehat{h \star y}$$

$$\widehat{z} = \widehat{h} \widehat{y}$$

$$\frac{\widehat{z}}{\widehat{h}} = \widehat{y}$$

valide lorsque $\widehat{h}[k_1, k_2] \neq 0$

Déconvolution par inversion

Modèle direct :

$$z = h \star y + n$$

avec $(n[k])_k$ variables aléatoires i.i.d Gaussiennes.

Solution par inversion :

$$\text{Rappel : } \hat{h}[k_1, k_2] = \sum_{\xi_1, \xi_2} h[\xi_1, \xi_2] e^{-\frac{i2\pi k_1 \xi_1}{N_1}} e^{-\frac{i2\pi k_2 \xi_2}{N_2}} .$$

$$z = h \star y$$

$$\hat{z} = \widehat{h \star y}$$

$$\hat{z} = \hat{h} \hat{y}$$

$$\frac{\hat{z}}{\hat{h}} = \hat{y}$$

valide lorsque $\hat{h}[k_1, k_2] \neq 0$

→ ceci est peut-être trop... simple!

Déconvolution par inversion

Modèle direct :

$$z = h \star y + n$$

avec $(n[k])_k$ variables aléatoires i.i.d Gaussiennes.

Solution par inversion :

$$\text{Rappel : } \hat{h}[k_1, k_2] = \sum_{\xi_1, \xi_2} h[\xi_1, \xi_2] e^{-\frac{i2\pi k_1 \xi_1}{N_1}} e^{-\frac{i2\pi k_2 \xi_2}{N_2}}.$$

$$z = h \star y$$

$$\hat{z} = \widehat{h \star y}$$

$$\hat{z} = \hat{h} \hat{y}$$

$$\frac{\hat{z}}{\hat{h}} = \hat{y}$$

valide lorsque $\hat{h}[k_1, k_2] \neq 0$

→ ceci est peut-être trop... simple! Ajoutons du bruit.

Déconvolution par inversion

Modèle direct :

$$z = h \star y + n$$

avec $(n[k])_k$ variables aléatoires i.i.d Gaussiennes.

Solution par inversion :

$$\text{Rappel : } \hat{h}[k_1, k_2] = \sum_{\xi_1, \xi_2} h[\xi_1, \xi_2] e^{-\frac{i2\pi k_1 \xi_1}{N_1}} e^{-\frac{i2\pi k_2 \xi_2}{N_2}} .$$

$$z = h \star y + n$$

$$\hat{z} = \widehat{h \star y} + \hat{n}$$

$$\hat{z} = \hat{h} \hat{y} + \hat{n}$$

$$\frac{\hat{z}}{\hat{h}} = \hat{y} + \frac{\hat{n}}{\hat{h}}$$

valide lorsque $\hat{h}[k_1, k_2] \neq 0$

→ ceci est peut-être trop... simple! Ajoutons du bruit.

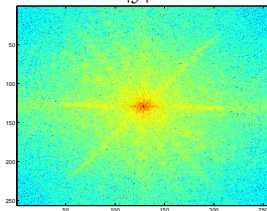
Les effets de la convolution

h est un “filtre uniforme” de longueur 4

Image originale y



$|\hat{y}|$



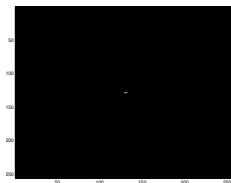
Les effets de la convolution

h est un "filtre uniforme" de longueur 4

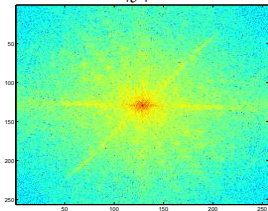
Image originale y



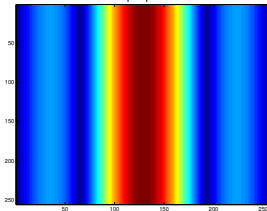
Filtre h



$|\hat{y}|$



$|\hat{h}|$



Les effets de la convolution

h est un "filtre uniforme" de longueur 4

y



$\star h$

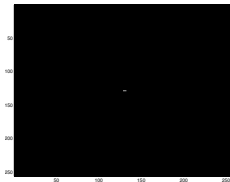
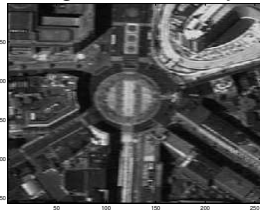
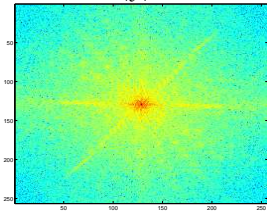


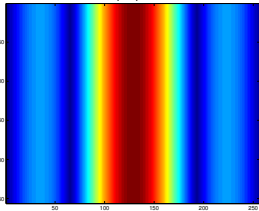
Image convolée $h \star y$



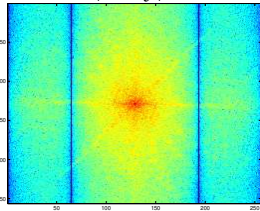
$|\hat{y}|$



$|\hat{h}|$



$|\hat{h} \times \hat{y}|$



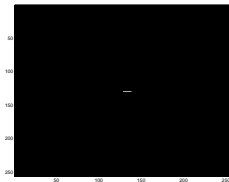
Les effets de la convolution

h est un "filtre médian" de longueur 10

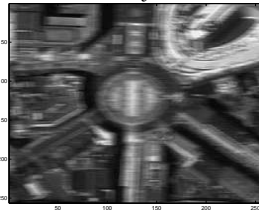
Image originale y



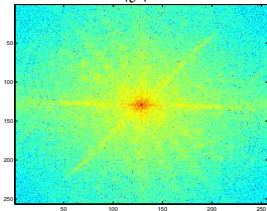
h



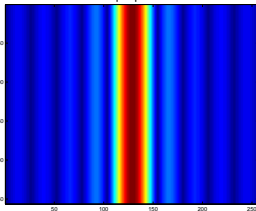
$h \star y$



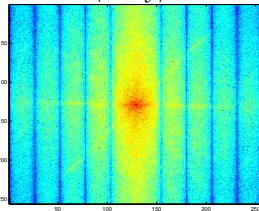
$|\hat{y}|$



$|\hat{h}|$



$|\hat{h} \times \hat{y}|$



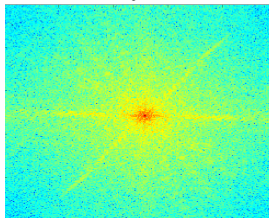
Les effets de la convolution

h est un "filtre Gaussien" ($\sigma = 10$)

Image originale y



$|\hat{y}|$



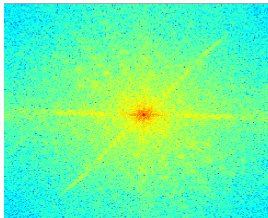
Les effets de la convolution

h est un "filtre Gaussien" ($\sigma = 10$)

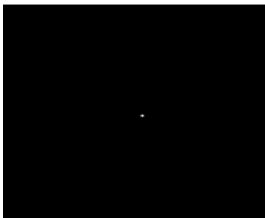
Image originale y



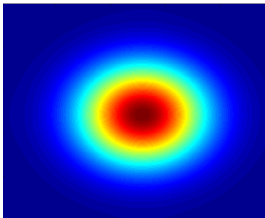
$|\hat{y}|$



Filtre h



$|\hat{h}|$



Les effets de la convolution

h est un "filtre Gaussien" ($\sigma = 10$)

y



$\star h$

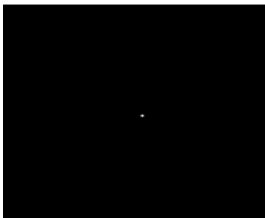
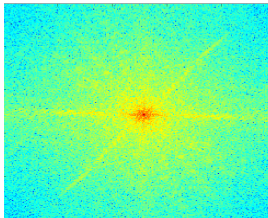


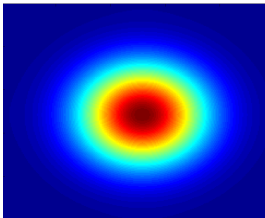
Image convolée $h \star y$



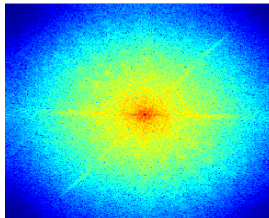
$|\hat{y}|$



$|\hat{h}|$



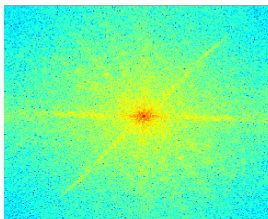
$|\hat{h} \times \hat{y}|$



Déconvolution

Inversion du filtre

Image originale



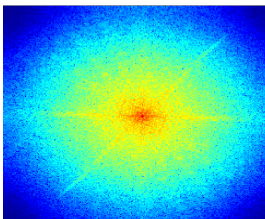
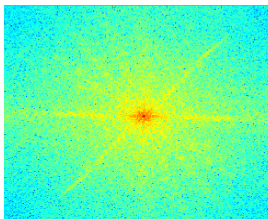
Déconvolution

Inversion du filtre

Image originale



Image convolée



Déconvolution

Inversion du filtre

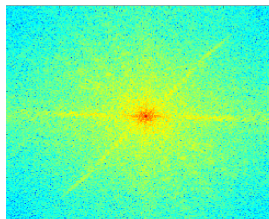
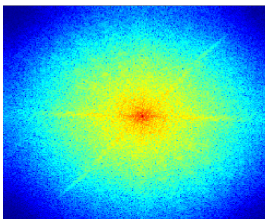
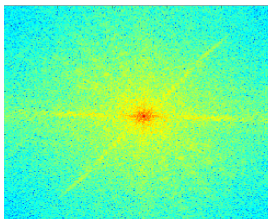
Image originale



Image convolée



Reconstruction



Déconvolution

Inversion du filtre : inconvénients !

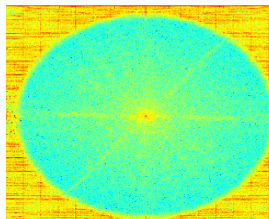
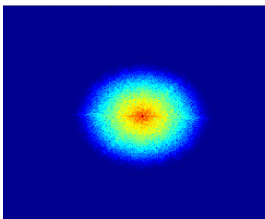
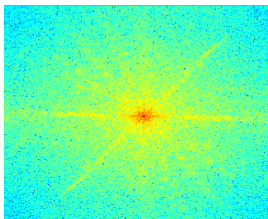
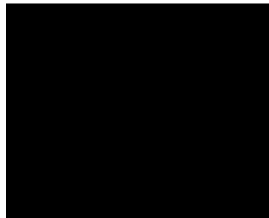
Image originale



Image convolée



Reconstruction



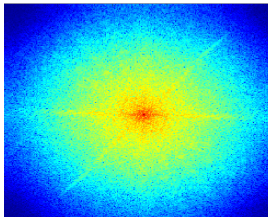
Déconvolution

Inversion du filtre : effet du bruit

$$h \star y$$



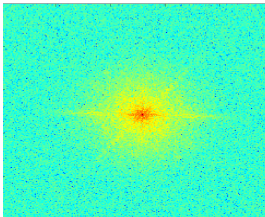
$$|\hat{h} \times \hat{y}|$$



$$h \star y + n, n \sim \mathcal{N}(0, 10)$$



$$|\hat{h} \times \hat{y} + \hat{n}|$$



Déconvolution

Inversion du filtre : effet du bruit

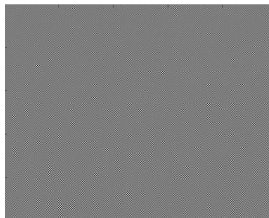
$h \star y$



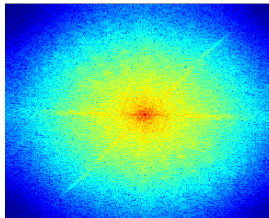
$h \star y + n, n \sim \mathcal{N}(0, 10)$



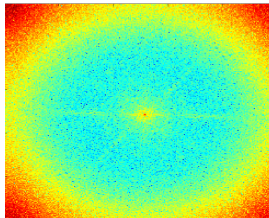
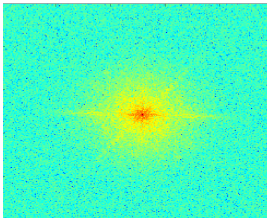
Inversion directe



$|\hat{h} \times \hat{y}|$



$|\hat{h} \times \hat{y} + \hat{n}|$



Inversion de l'opérateur L

L^{-1} peut exister mais

- ▶ être difficile à calculer
- ▶ amener à des erreurs numériques...

Inversion de l'opérateur L

L^{-1} peut exister mais

- ▶ être difficile à calculer
- ▶ amener à des erreurs numériques...

Conditionnement

Pour contrôler ces instabilités numériques on calcule le **conditionnement** de L

$$k(L) = \|L\|_2 \|L^{-1}\|_2 = \frac{\sigma_{\max}(L)}{\sigma_{\min}(L)}$$

en effet si

Inversion de l'opérateur L

L^{-1} peut exister mais

- ▶ être difficile à calculer
- ▶ amener à des erreurs numériques...

Conditionnement

Pour contrôler ces instabilités numériques on calcule le **conditionnement** de L

$$k(L) = \|L\|_2 \|L^{-1}\|_2 = \frac{\sigma_{\max}(L)}{\sigma_{\min}(L)}$$

en effet si

$$z = Ly$$

Inversion de l'opérateur L

L^{-1} peut exister mais

- ▶ être difficile à calculer
- ▶ amener à des erreurs numériques...

Conditionnement

Pour contrôler ces instabilités numériques on calcule le **conditionnement** de L

$$k(L) = \|L\|_2 \|L^{-1}\|_2 = \frac{\sigma_{\max}(L)}{\sigma_{\min}(L)}$$

en effet si

$$\begin{aligned} z &= Ly \\ z + \delta z &= L(y + \delta y) \end{aligned}$$

Inversion de l'opérateur L

L^{-1} peut exister mais

- ▶ être difficile à calculer
- ▶ amener à des erreurs numériques...

Conditionnement

Pour contrôler ces instabilités numériques on calcule le **conditionnement** de L

$$k(L) = \|L\|_2 \|L^{-1}\|_2 = \frac{\sigma_{\max}(L)}{\sigma_{\min}(L)}$$

en effet si

$$\begin{aligned} z &= Ly \\ z + \delta z &= L(y + \delta y) \quad \text{avec} \quad \delta z = L\delta y \end{aligned}$$

Inversion de l'opérateur L

L^{-1} peut exister mais

- ▶ être difficile à calculer
- ▶ amener à des erreurs numériques...

Conditionnement

Pour contrôler ces instabilités numériques on calcule le **conditionnement** de L

$$k(L) = \|L\|_2 \|L^{-1}\|_2 = \frac{\sigma_{\max}(L)}{\sigma_{\min}(L)}$$

en effet si

$$\begin{aligned} z &= Ly \\ z + \delta z &= L(y + \delta y) \quad \text{avec} \quad \delta z = L\delta y \\ \|\delta y\|_2 &\leq \|L^{-1}\| \|\delta z\| \end{aligned}$$

Inversion de l'opérateur L

L^{-1} peut exister mais

- ▶ être difficile à calculer
- ▶ amener à des erreurs numériques...

Conditionnement

Pour contrôler ces instabilités numériques on calcule le **conditionnement** de L

$$k(L) = \|L\|_2 \|L^{-1}\|_2 = \frac{\sigma_{\max}(L)}{\sigma_{\min}(L)}$$

en effet si

$$\begin{aligned} z &= Ly \\ z + \delta z &= L(y + \delta y) \quad \text{avec} \quad \delta z = L\delta y \\ \|\delta y\|_2 &\leq \|L^{-1}\| \|\delta z\| \\ \frac{1}{\|y\|} &\leq \|L\| \frac{1}{\|z\|} \end{aligned}$$

Inversion de l'opérateur L

L^{-1} peut exister mais

- ▶ être difficile à calculer
- ▶ amener à des erreurs numériques...

Conditionnement

Pour contrôler ces instabilités numériques on calcule le **conditionnement** de L

$$k(L) = \|L\|_2 \|L^{-1}\|_2 = \frac{\sigma_{\max}(L)}{\sigma_{\min}(L)}$$

en effet si

$$\begin{aligned} z &= Ly \\ z + \delta z &= L(y + \delta y) \quad \text{avec} \quad \delta z = L\delta y \\ \|\delta y\|_2 &\leq \|L^{-1}\| \|\delta z\| \\ \frac{1}{\|y\|} &\leq \|L\| \frac{1}{\|z\|} \\ \frac{\|\delta y\|}{\|y\|} &\leq k(L) \frac{\|\delta z\|}{\|z\|} \end{aligned}$$

Pseudo-inversion de l'opérateur L

Supposons que L^{-1} n'existe pas (par ex. $\ker L \neq \{0\}$) et résolvons le problème aux moindres carrés suivant

$$\tilde{y} = \arg \min_y \frac{1}{2} \| Ly - z \|_2^2$$

↔ la solution n'est pas unique, en revanche il existe une unique solution de norme minimale...

Pseudo-inversion de l'opérateur L

Supposons que L^{-1} n'existe pas (par ex. $\ker L \neq \{0\}$) et résolvons le problème aux moindres carrés suivant

$$\tilde{y} = \arg \min_y \frac{1}{2} \|Ly - z\|_2^2$$

↔ la solution n'est pas unique, en revanche il existe une unique solution de norme minimale...

- ▶ En effet tous les minimiseurs \tilde{y} de $F(y) = \frac{1}{2} \|Ly - z\|_2^2$ vérifient

$$\nabla F(\tilde{y}) = 0 \Leftrightarrow L^*L\tilde{y} = L^*z$$

Pseudo-inversion de l'opérateur L

Supposons que L^{-1} n'existe pas (par ex. $\ker L \neq \{0\}$) et résolvons le problème aux moindres carrés suivant

$$\tilde{y} = \arg \min_y \frac{1}{2} \|Ly - z\|_2^2$$

↔ la solution n'est pas unique, en revanche il existe une unique solution de norme minimale...

- ▶ En effet tous les minimiseurs \tilde{y} de $F(y) = \frac{1}{2} \|Ly - z\|_2^2$ vérifient

$$\nabla F(\tilde{y}) = 0 \Leftrightarrow L^*L\tilde{y} = L^*z$$

- ▶ $\tilde{y} = \underbrace{y_0}_{\in \text{Ker}(L)} + \underbrace{y_\perp}_{\in \text{Ker}(L)^\perp}$ implique $L^*Ly_\perp = L^*L\tilde{y} = L^*z$,

Pseudo-inversion de l'opérateur L

Supposons que L^{-1} n'existe pas (par ex. $\ker L \neq \{0\}$) et résolvons le problème aux moindres carrés suivant

$$\tilde{y} = \arg \min_y \frac{1}{2} \|Ly - z\|_2^2$$

↔ la solution n'est pas unique, en revanche il existe une unique solution de norme minimale...

- ▶ En effet tous les minimiseurs \tilde{y} de $F(y) = \frac{1}{2} \|Ly - z\|_2^2$ vérifient

$$\nabla F(\tilde{y}) = 0 \Leftrightarrow L^*L\tilde{y} = L^*z$$

- ▶ $\tilde{y} = \underbrace{y_0}_{\in \ker(L)} + \underbrace{y_{\perp}}_{\in \ker(L)^{\perp}}$ implique $L^*Ly_{\perp} = L^*L\tilde{y} = L^*z$,

- ▶ ainsi une solution est de norme minimale si $y_0 = 0$ et on a

$$\begin{aligned} L^*Ly_{\perp} &= L^*z \\ ((L^*L)_{\ker(L)^{\perp}})^{-1} L^*Ly_{\perp} &= \underbrace{((L^*L)_{\ker(L)^{\perp}})^{-1} L^*z}_{\text{pseudoinverse}} \end{aligned}$$

$$y_{\perp} = L^{\dagger}z$$

Pseudo-inversion de l'opérateur

L'algorithme de Landweber pour calculer L^\dagger :

- ▶ Initialisation : $y^0 = 0$
- ▶ Itération : $y^{k+1} = y^k - \omega L^*(Ly^k - z)$

On a :

$$0 < \omega < \frac{2}{\sigma_1(L)^2} \Rightarrow \lim_{k \rightarrow \infty} y^k = L^\dagger z$$

Pseudo-inversion de l'opérateur

Malheureusement...

La pseudo-inverse a les mêmes problèmes que L^{-1} :

- ▶ inversion numérique,
- ▶ conditionnement de L^\dagger ,
- ▶ problèmes avec le bruit.

Voyons la **SVD tronquée**.

Inversion par SVD tronquée

“Singular Value Decomposition” ou décomposition en valeurs singulières.

SVD de L : $L = U\Delta V^*$

- ▶ $L \in \mathbb{R}^{M \times N}$
- ▶ $U = (u_1, \dots, u_M)$ est unitaire ($U^*U = \text{Id}_M$),
- ▶ $V = (v_1, \dots, v_N)$ est unitaire ($V^*V = \text{Id}_N$).
- ▶ $\Delta = \text{diag}(d_1, \dots, d_k, 0, \dots, 0)$ est diagonale.
- ▶ d_1, \dots, d_k sont les valeurs singulières de L
- ▶ $|d_1| \geq |d_2|, \dots, \geq |d_k| > 0$.

Inversion par SVD tronquée

“Singular Value Decomposition” ou décomposition en valeurs singulières.

SVD de L : $L = U\Delta V^*$

- ▶ $L \in \mathbb{R}^{M \times N}$
- ▶ $U = (u_1, \dots, u_M)$ est unitaire ($U^*U = \text{Id}_M$),
- ▶ $V = (v_1, \dots, v_N)$ est unitaire ($V^*V = \text{Id}_N$).
- ▶ $\Delta = \text{diag}(d_1, \dots, d_k, 0, \dots, 0)$ est diagonale.
- ▶ d_1, \dots, d_k sont les valeurs singulières de L
- ▶ $|d_1| \geq |d_2|, \dots, \geq |d_k| > 0$.

Propriétés de la SVD

- ▶ $L(v_i) = d_i u_i$ si $i = 1 \dots k$,
- ▶ $L(v_i) = 0$ si $i > k$.
- ▶ $L(\mathbf{y}) = \sum_{i=1}^k d_i \langle \mathbf{y}, v_i \rangle u_i$.
- ▶ Conditionnement : $|d_1|/|d_k|$.
- ▶ Pseudo-inverse : $L^\dagger = V\Delta^{-1}U^*$ avec $\Delta^{-1} = \text{diag}(1/d_1, \dots, 1/d_k, 0, \dots, 0)$.

Inversion par SVD tronquée

Propriétés

- ▶ Pseudo-inverse : $L^\dagger = V\Delta^{-1}U^*$ avec $\Delta^{-1} = \text{diag}(1/d_1, \dots, 1/d_k, 0, \dots, 0)$.
- ▶ Conditionnement de L^\dagger : $|d_1|/|d_k|$.

SVD tronquée

Fixons $p < k$:

$$L_{svd}^{-1} = V\Delta_p^{-1}U^*$$

- ▶ $\Delta_p^{-1} = \text{diag}(1/d_1, \dots, 1/d_p, 0, \dots, 0)$

Inversion par SVD tronquée

Propriétés

- ▶ Pseudo-inverse : $L^\dagger = V\Delta^{-1}U^*$ avec $\Delta^{-1} = \text{diag}(1/d_1, \dots, 1/d_k, 0, \dots, 0)$.
- ▶ Conditionnement de L^\dagger : $|d_1|/|d_k|$.

SVD tronquée

Fixons $p < k$:

$$L_{svd}^{-1} = V\Delta_p^{-1}U^*$$

- ▶ $\Delta_p^{-1} = \text{diag}(1/d_1, \dots, 1/d_p, 0, \dots, 0)$
- ▶ Conditionnement : $|d_1|/|d_p| \rightarrow$ plus stable que la (pseudo-)inversion.
- ▶ On ne peut pas privilégier de “bonnes propriétés” de l’estimée y .

Estimateur de Wiener

Définition de l'estimateur de Wiener

- ▶ Les signaux (y, z) sont des réalisations de processus aléatoires stationnaires.
- ▶ L'estimateur de Wiener de y à partir des observations z est l'**estimateur linéaire** $\tilde{y} = Dz$ qui **minimise le risque**

$$\mathbb{E}\{\|\tilde{y} - y\|_2^2\}.$$

- ▶ D linéaire signifie

$$\forall n, \tilde{y}[n] = \sum_k D[n, k]z[k].$$

Estimateur de Wiener pour le débruitage de vecteurs Gaussiens

Modèle

- ▶ $z = y + w$
- ▶ Les signaux finis y , w et z sont des réalisations de vecteurs aléatoires Y , Z et W .
- ▶ Hypothèse Gaussienne : $Y \sim \mathcal{N}(0, \Sigma_Y)$, $W \sim \mathcal{N}(0, \Sigma_W)$.

Théorème

L'estimateur de Wiener i.e. l'estimateur linéaire $\tilde{y} = Dz$ qui minimise

$$\mathbb{E}\{\|\tilde{y} - y\|_2^2\}$$

est

$$\tilde{y} = \Sigma_Y(\Sigma_Y + \Sigma_W)^{-1}z.$$

Estimateur de Wiener pour le débruitage

Modèle

- ▶ $z = y + w$
- ▶ Les signaux finis y , w et z sont des réalisations de vecteurs aléatoires Y , Z et W stationnaires au sens large.
- ▶ Covariances : $\mathbb{E}\{W[n]W[k]\} = \Sigma_W[n - k]$, $\mathbb{E}\{Y[n]Y[k]\} = \Sigma_Y[n - k]$.
- ▶ Puissance spectrale : $\widehat{\Sigma}_Y$, $\widehat{\Sigma}_W$.

Théorème

L'estimateur de Wiener i.e. l'estimateur linéaire $\tilde{y} = Dz$ qui minimise

$$\mathbb{E}\{\|\tilde{y} - y\|_2^2\}$$

est décrit dans la base de Fourier par

$$\widehat{\tilde{y}}[m] = \frac{\widehat{\Sigma}_Y[m]}{\widehat{\Sigma}_Y[m] + \widehat{\Sigma}_W[m]} \widehat{z}[m]$$

Estimateur de Wiener pour la déconvolution

Modèle

- ▶ $z = h \star y + w$
- ▶ Les signaux finis y, w et z sont des réalisations de vecteurs aléatoires Y, Z et W stationnaires au sens large.
- ▶ Covariances : $\mathbb{E}\{W[n]W[k]\} = \Sigma_W[n - k]$, $\mathbb{E}\{Y[n]Y[k]\} = \Sigma_Y[n - k]$.
- ▶ Puissance spectrale : $\widehat{\Sigma}_Y, \widehat{\Sigma}_W$.

Théorème

L'estimateur de Wiener i.e. l'estimateur linéaire $\tilde{y} = Dz$ qui minimise

$$\mathbb{E}\{\|\tilde{y} - y\|_2^2\}$$

est décrit dans la base de Fourier par

$$\widehat{\tilde{y}}[m] = \frac{\widehat{\Sigma}_Y[m]}{|\widehat{h}[m]|^2 \widehat{\Sigma}_Y[m] + \widehat{\Sigma}_W[m]} \overline{\widehat{h}[m]} \widehat{z}[m]$$

Information sur la solution cible

- ▶ Le respect des données ne suffit pas ...
- ▶ Solution : ajouter de l'information sur la solution recherchée.

Ajout d'information a priori

- ▶ sous forme de contrainte

$$\tilde{y} = \arg \min_y \frac{1}{2} \| Ly - z \|_2^2 \text{ s.c. } y \in \mathcal{Y}$$

- ▶ sous forme de paramétrisation

$$\tilde{\theta} = \arg \min_{\theta} \frac{1}{2} \| Ly(\theta) - z \|_2^2$$

- ▶ sous forme de **régularisation**

$$\tilde{y}_{\lambda} = \arg \min_y \frac{1}{2} \| Ly - z \|_2^2 + \lambda \Omega(y).$$

Plan de la séance

INTRODUCTION ET EXEMPLES

- Introduction
- Exemples

L'INVERSION, UN PROBLÈME DIFFICILE

- Problèmes mal posés
- Résolution par (pseudo-)inversion
- SVD tronquée
- Estimateur de Wiener

RÉGULARISATION

- Régularisation de Tikhonov
- La régularisation d'un point de vue variationnel et statistique
- MAP

CONCLUSIONS

Régularisation du problème

Méthode de régularisation

Trouver $\lambda \geq 0$ et $\tilde{y}_\lambda = \arg \min_y \frac{1}{2} \|Ly - z\|_2^2 + \lambda \Omega(y)$

Régularisation du problème

Méthode de régularisation

Trouver $\lambda \geq 0$ et $\tilde{y}_\lambda = \arg \min_y \frac{1}{2} \|Ly - z\|_2^2 + \lambda \Omega(y)$

- ▶ $\frac{1}{2} \|Ly - z\|_2^2$: **terme de fidélité**.
- ▶ $\Omega(y)$: pénalisation ou **régularisation** choisi comme un a priori sur la solution.
- ▶ λ : paramètre de régularisation.

Régularisation du problème

Méthode de régularisation

Trouver $\lambda \geq 0$ et $\tilde{y}_\lambda = \arg \min_y \frac{1}{2} \|Ly - z\|_2^2 + \lambda \Omega(y)$

- ▶ $\frac{1}{2} \|Ly - z\|_2^2$: **terme de fidélité**.
- ▶ $\Omega(y)$: pénalisation ou **régularisation** choisi comme un a priori sur la solution.
- ▶ λ : paramètre de régularisation.

Cas particuliers

- ▶ $\lambda = 0$: problème des moindres carrés étudié précédemment.

Régularisation du problème

Méthode de régularisation

Trouver $\lambda \geq 0$ et $\tilde{y}_\lambda = \arg \min_y \frac{1}{2} \|Ly - z\|_2^2 + \lambda \Omega(y)$

- ▶ $\frac{1}{2} \|Ly - z\|_2^2$: **terme de fidélité**.
- ▶ $\Omega(y)$: pénalisation ou **régularisation** choisi comme un a priori sur la solution.
- ▶ λ : paramètre de régularisation.

Cas particuliers

- ▶ $\lambda = 0$: problème des moindres carrés étudié précédemment.
- ▶ $\Omega(y) = \|y\|^2$: **régularisation de Tikhonov**.

Régularisation du problème

Méthode de régularisation

Trouver $\lambda \geq 0$ et $\tilde{y}_\lambda = \arg \min_y \frac{1}{2} \|Ly - z\|_2^2 + \lambda \Omega(y)$

- ▶ $\frac{1}{2} \|Ly - z\|_2^2$: **terme de fidélité**.
- ▶ $\Omega(y)$: pénalisation ou **régularisation** choisi comme un a priori sur la solution.
- ▶ λ : paramètre de régularisation.

Cas particuliers

- ▶ $\lambda = 0$: problème des moindres carrés étudié précédemment.
- ▶ $\Omega(y) = \|y\|^2$: **régularisation de Tikhonov**.
↪ méthode similaire pour trouver l'unique solution

$$\tilde{y}_\lambda = (L^*L + 2\lambda I)^{-1}L^*z.$$

Régularisation du problème

Méthode de régularisation

Trouver $\lambda \geq 0$ et $\tilde{y}_\lambda = \arg \min_y \frac{1}{2} \|Ly - z\|_2^2 + \lambda \Omega(y)$

- ▶ $\frac{1}{2} \|Ly - z\|_2^2$: **terme de fidélité**.
- ▶ $\Omega(y)$: pénalisation ou **régularisation** choisi comme un a priori sur la solution.
- ▶ λ : paramètre de régularisation.

Cas particuliers

- ▶ $\lambda = 0$: problème des moindres carrés étudié précédemment.
- ▶ $\Omega(y) = \|y\|^2$: **régularisation de Tikhonov**.
↪ méthode similaire pour trouver l'unique solution

$$\tilde{y}_\lambda = (L^*L + 2\lambda I)^{-1}L^*z.$$

Application à la déconvolution :

$$\widehat{y}_\lambda = \frac{\widehat{h}}{|\widehat{h}|^2 + 2\lambda} \widehat{z}.$$

Régularisation de Tikhonov

Image originale



Image floue et bruitée



Régularisation de Tikhonov

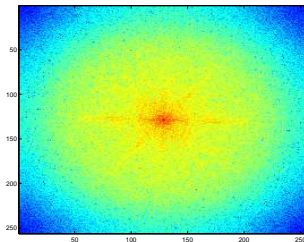
Image originale



Image floue et bruitée



Reconstruction $\lambda = 0.05$



Régularisation de Tikhonov

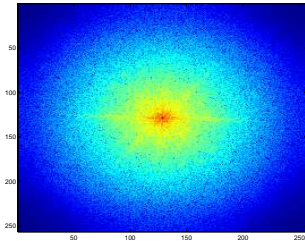
Image originale



Image floue et bruitée



Reconstruction $\lambda = 5$



La régularisation d'un point de vue variationnel et statistique

Méthode de régularisation

Pour le modèle direct $z = \mathcal{D}_\alpha(Ly)$, trouver

$$\tilde{y}_\lambda = \arg \min_y \frac{1}{2} \|Ly - z\|_2^2 + \lambda \Omega(y)$$

- ▶ $\frac{1}{2} \|Ly - z\|_2^2$: terme de fidélité (bruit, \mathcal{D}_α)
- ▶ $\lambda \Omega(y)$: a priori sur y .

La régularisation d'un point de vue variationnel et statistique

Méthode de régularisation

Pour le modèle direct $z = \mathcal{D}_\alpha(Ly)$, trouver

$$\tilde{y} = \arg \min_y f_1(y) + f_2(y)$$

- ▶ $f_1(y)$: terme de fidélité (bruit, \mathcal{D}_α)
- ▶ $f_2(y)$: a priori sur y .

La régularisation d'un point de vue variationnel et statistique

Méthode de régularisation

Pour le modèle direct $z = \mathcal{D}_\alpha(Ly)$, trouver

$$\tilde{y} = \arg \min_y f_1(y) + f_2(y)$$

- ▶ $f_1(y)$: terme de fidélité (bruit, \mathcal{D}_α)
- ▶ $f_2(y)$: a priori sur y .

Exemple (Tikhonov : un point de vue variationnel)

Chercher un compromis entre 2 espaces fonctionnels

- ▶ $f_1(y) = \frac{1}{2} \| Ly - z \|_2^2$: la distance entre les images est mesurée au sens l_2 ,
- ▶ $f_2(y) = \lambda \| y \|_2^2$: les y solutions ont une petite norme l_2 .

La régularisation d'un point de vue variationnel et statistique

Méthode de régularisation

Pour le modèle direct $z = \mathcal{D}_\alpha(Ly)$, trouver

$$\tilde{y} = \arg \min_y f_1(y) + f_2(y)$$

- ▶ $f_1(y)$: terme de fidélité (bruit, \mathcal{D}_α)
- ▶ $f_2(y)$: a priori sur y .

Exemple (Tikhonov : un point de vue variationnel)

Chercher un compromis entre 2 espaces fonctionnels

- ▶ $f_1(y) = \frac{1}{2} \| Ly - z \|_2^2$: la distance entre les images est mesurée au sens l_2 ,
- ▶ $f_2(y) = \lambda \| y \|_2^2$: les y solutions ont une petite norme l_2 .

Exemple (Tikhonov : un point de vue statistique)

Estimation au sens du Maximum a Posteriori quand

- ▶ $f_1(y) = \frac{1}{2} \| Ly - z \|_2^2$: le bruit est blanc gaussien additif,
- ▶ $f_2(y) = \lambda \| y \|_2^2$: y est un processus blanc gaussien.

Estimateur du Maximum A Posteriori (MAP)

- ▶ y et z sont des réalisations de vecteurs aléatoires Y et Z .
- ▶ Trouver le signal inconnu noté ici y grâce à l'estimateur du **Maximum A Posteriori** (MAP), c'est trouver y qui maximise la distribution a posteriori des pixels de l'image $\mu_{Y|Z=z}(y)$.
- ▶ On utilise la formule de Bayes : $\mu_{Y|Z=z}(y)\mu_Z(z) = \mu_{Z|Y=y}(z)\mu_Y(y)$ pour écrire le MAP

$$\hat{y} \in \underset{y \in \mathbb{R}^N}{\text{Argmax}} \mu_{Y|Z=z}(y)$$

$$\Leftrightarrow \hat{y} \in \underset{y \in \mathbb{R}^N}{\text{Argmax}} \mu_{Z|Y=y}(z)\mu_Y(y)$$

$$\Leftrightarrow \hat{y} \in \underset{y \in \mathbb{R}^N}{\text{Argmin}} \{ -\log(\mu_{Z|Y=y}(z)) - \log(\mu_Y(y)) \}$$

Estimateur du Maximum A Posteriori (MAP)

- ▶ y et z sont des réalisations de vecteurs aléatoires Y et Z .
- ▶ Trouver le signal inconnu noté ici y grâce à l'estimateur du **Maximum A Posteriori** (MAP), c'est trouver y qui maximise la distribution a posteriori des pixels de l'image $\mu_{Y|Z=z}(y)$.
- ▶ On utilise la formule de Bayes : $\mu_{Y|Z=z}(y)\mu_Z(z) = \mu_{Z|Y=y}(z)\mu_Y(y)$ pour écrire le MAP

$$\hat{y} \in \underset{y \in \mathbb{R}^N}{\text{Argmax}} \mu_{Y|Z=z}(y)$$

$$\Leftrightarrow \hat{y} \in \underset{y \in \mathbb{R}^N}{\text{Argmax}} \mu_{Z|Y=y}(z)\mu_Y(y)$$

$$\Leftrightarrow \hat{y} \in \underset{y \in \mathbb{R}^N}{\text{Argmin}} \{ -\log(\mu_{Z|Y=y}(z)) - \log(\mu_Y(y)) \}$$

Lien avec une approche variationnelle

- $-\log(\mu_{Z|Y=y}(z))$: terme de fidélité aux données
- $-\log(\mu_Y(y))$: terme d'*a priori*

Formulation générale du problème

Problème de minimization

$$\min_{y \in \mathbb{R}^N} \underbrace{f_1(y)}_{\text{Terme de fidélité aux données}} + \underbrace{f_2(y)}_{\text{Terme d'a priori}}$$

Question : comment déterminer f_1 et f_2 ?

Plan de la séance

INTRODUCTION ET EXEMPLES

- Introduction
- Exemples

L'INVERSION, UN PROBLÈME DIFFICILE

- Problèmes mal posés
- Résolution par (pseudo-)inversion
- SVD tronquée
- Estimateur de Wiener

RÉGULARISATION

- Régularisation de Tikhonov
- La régularisation d'un point de vue variationnel et statistique
- MAP

CONCLUSIONS

Conclusions

Résoudre un problème inverse, c'est dur

à cause

- ▶ de l'opérateur,
- ▶ des dégradations.

Pour améliorer les résultats d'inversion, on prend en compte des a-priori sur les objets, par exemple la régularisation.

On a vu

des approches statistiques, variationnelles.

On verra demain

- ▶ des régularisations,
- ▶ des algorithmes.