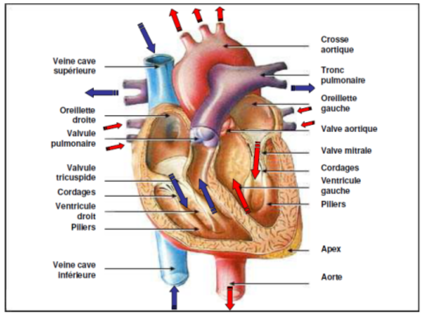
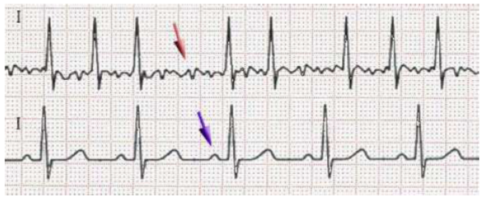
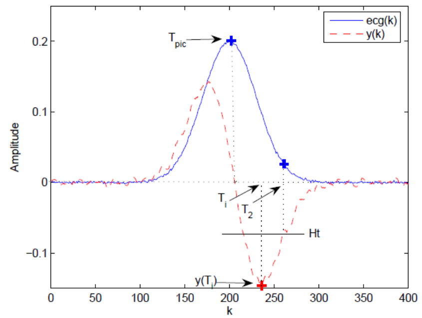
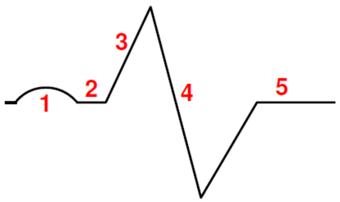
Caractérisation des signaux physiologiques

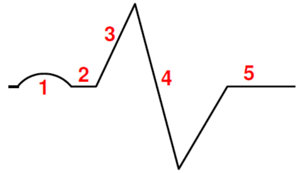
5. Filtrage adaptatif







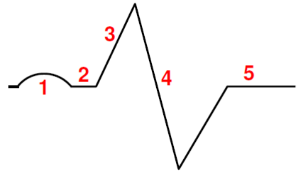




Plan

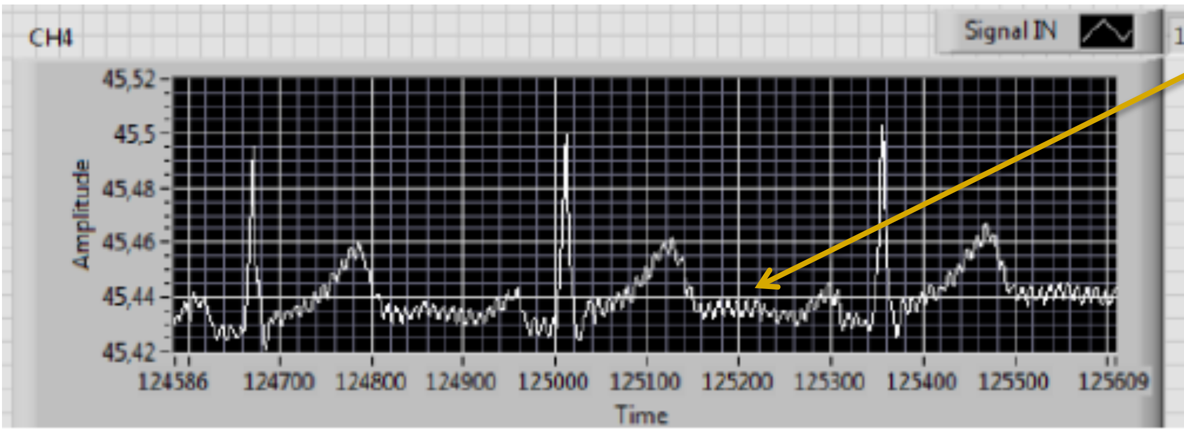
- Introduction

Introduction



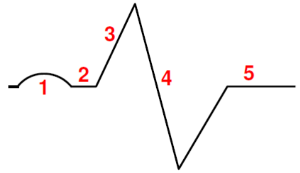
- **Contexte**
  - Les signaux physiologiques sont des signaux à composante aléatoire, l'acquisition étant perturbée par différents types de bruit

  - **Exemples :**



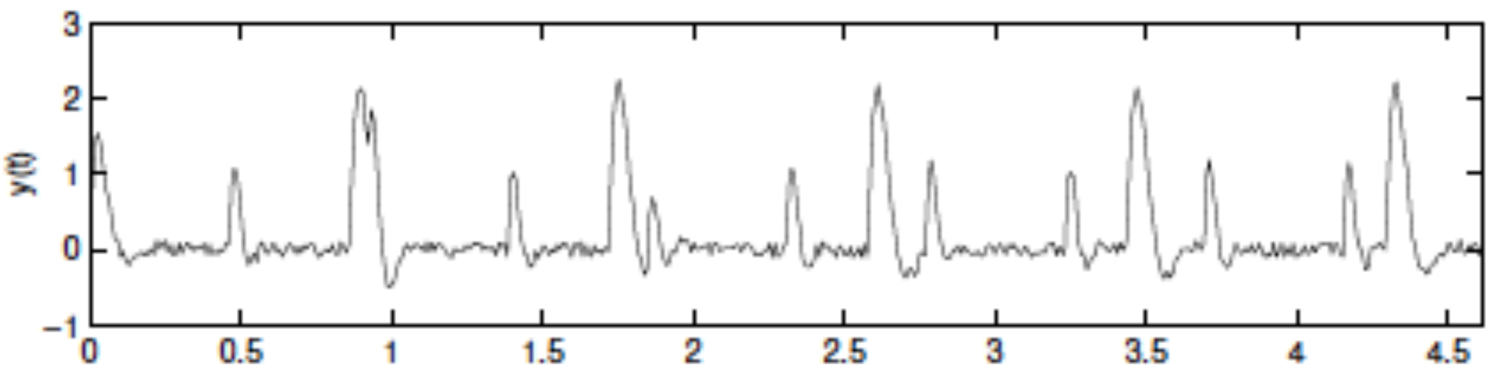
Bruit d'acquisition

Le bruit à 50Hz sur un ECG



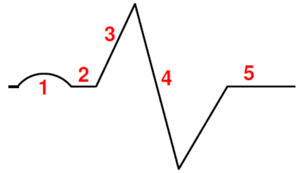
Introduction

- **Contexte**
  - Les signaux physiologiques sont des signaux à composante aléatoire, l'acquisition étant perturbée par différents types de bruit
  - **Exemples :**



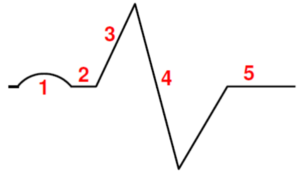
ECG fœtal perturbé par celui de la maman

Introduction



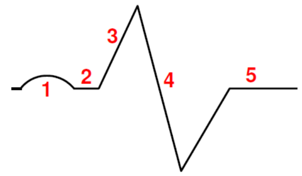
- **Contexte**
  - Le filtrage classique atteint ses limites pour ce type de signaux (superposition des spectres)
  - L'idée est alors de définir des filtres dont **la réponse impulsionnelle peut s'adapter aux évolutions des caractéristiques statistiques du signal étudié** (moyenne, variance, corrélation etc.) au cours du temps afin d'extraire la composante d'intérêt

Introduction



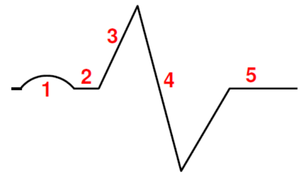
- **Filtrage adaptatif ?**
  - Le filtrage adaptatif est basé sur la recherche de **paramètres optimaux** par **minimisation d'un critère de performance.**
  - Les signaux étudiés ici étant à composante aléatoire, il est nécessaire d'avoir en tête les estimateurs associés
  - Ce type de filtrage faisant appel à une problématique de minimisation d'un critère par rapport à un ensemble d'observation, le principe de régression linéaire sera également revu…

Introduction



- **Filtrage adaptatif ?**
  - …pour enfin s’attaquer au filtre de Wiener…
  - … et à la notion de filtrage adaptatif.

Plan



- Introduction
- Rappels sur l'estimation statistique

Estimation statistique

- **Rappels**
  - La moyenne d'une variable aléatoire ***X*** caractérisée par une ddp ***p(x) :***

$$\mu_X = E[X] = \int_{-\infty}^{+\infty} x p(x)\, dx$$

  - La variance de cette même variable aléatoire

$$\sigma_X^2 = E\left[\left(X-\mu_X\right)^2\right] = \int_{-\infty}^{+\infty} \left(x-\mu_X\right)^2 p(x)\, dx$$

$$\sigma_X^2 = E\left[X^2\right] - E[X]^2$$

Estimation statistique

- **Rappels**
  - Corrélation de deux variables aléatoires ***X** et **Y*** de lois de densité de probabilité conjointe ***p(x,y) :***

$$R_{XY} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyp(x,y)\,dx\,dy$$

  - La covariance associée

$$C_{XY} = \sigma_{XY} = E\left[\left(X-\mu_X\right)\left(Y-\mu_Y\right)\right]$$

$$C_{XY} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(x-\mu_X\right)\left(y-\mu_Y\right)p(x,y)\,dx\,dy$$

Estimation statistique

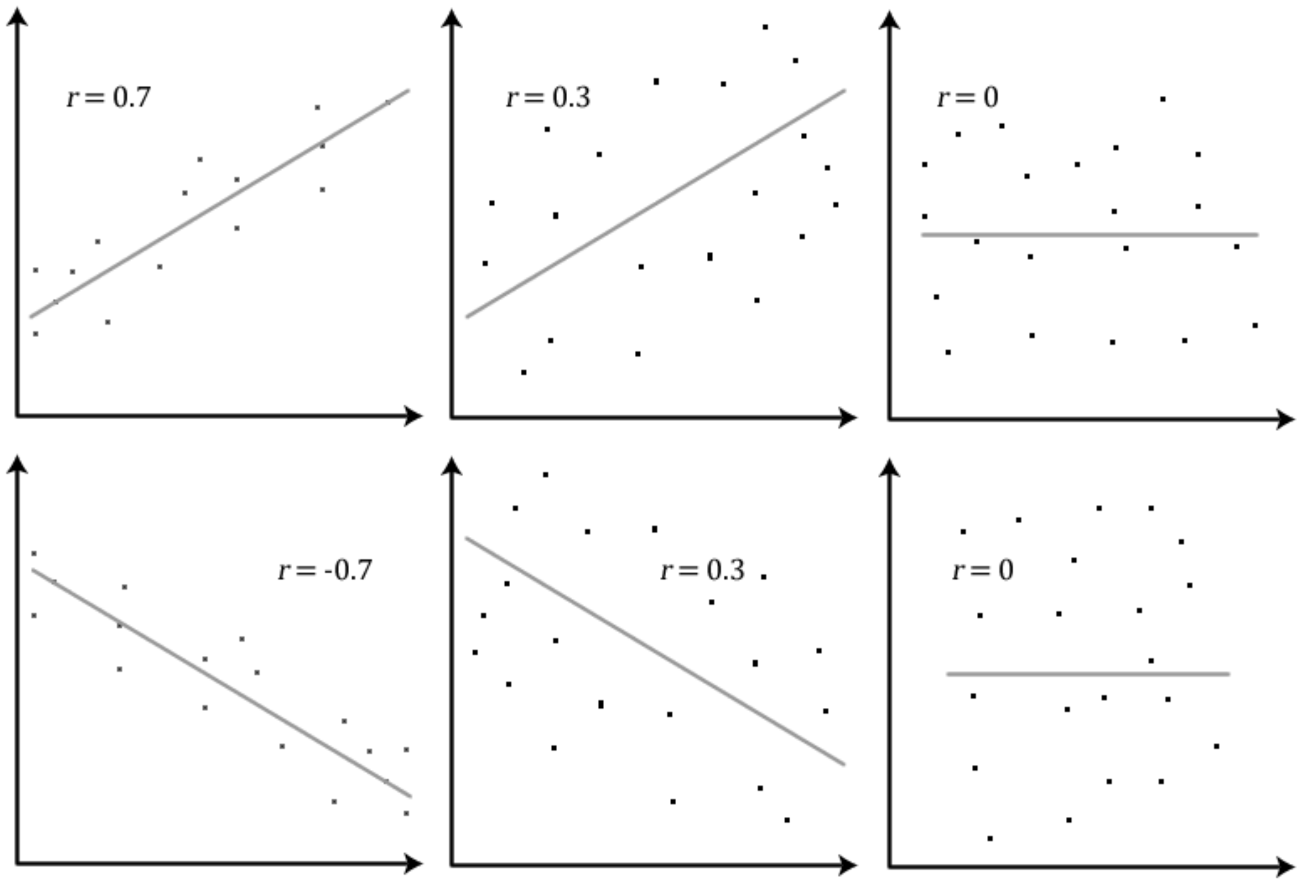
- **Rappels**
  - Coefficient de **corrélation linéaire**

$$r_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y} \in [-1,1]$$

Il s'agit d'une « quantification » de l'interdépendance de 2 *v.a*

Estimation statistique

- **Rappels**
  - Coefficient de **corrélation linéaire**



Plan

- Introduction
- Rappels sur l'estimation statistique
- Régression linéaire

Régression linéaire

Principe de la régression

- On souhaite expliquer les variations d’une variable **Y** à partir des valeurs observées pour la variable ***x***.

- Le problème n’est pas symétrique : les 2 variables n’ont pas le même statut :

Y	variable à expliquer ou réponse, **supposée aléatoire**
x	variable explicative ou covariable ou régresseur, supposée fixe

Régression linéaire

- **Principe de la régression**
  - **Modèle général**

$$Y = f(x) + E$$

Fonction associée à la régression

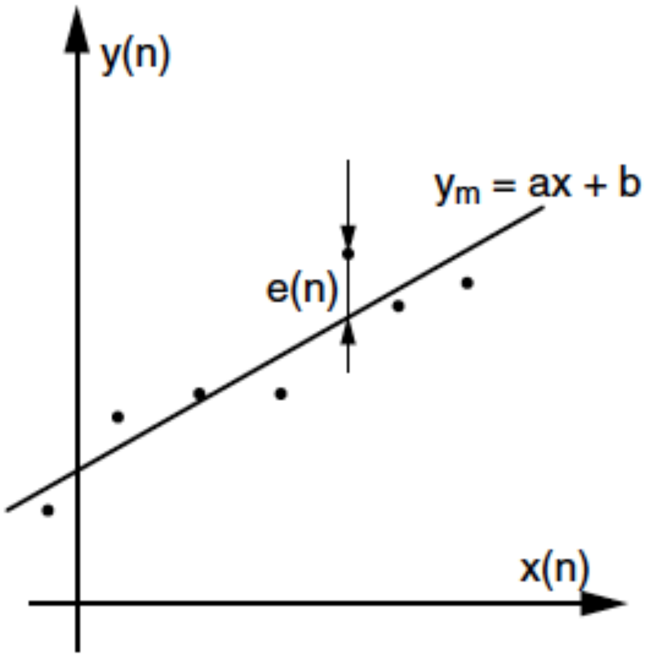
Terme résiduel aléatoire, ou erreur

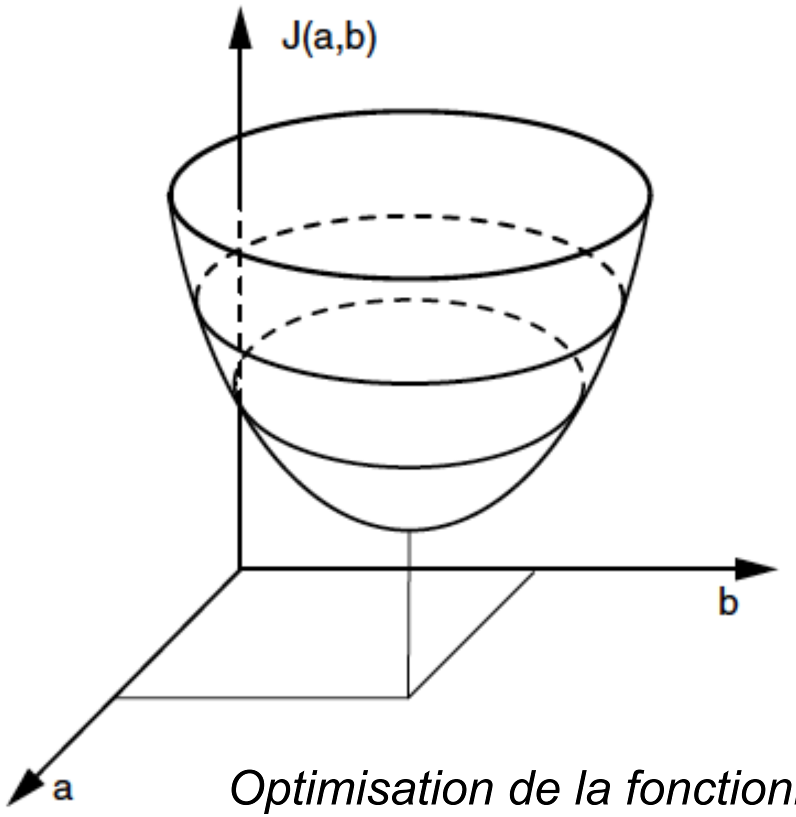
Régression linéaire

- **Principe**
  - La régression **linéaire** consiste en la recherche de **la droite** passant **au mieux parmi un ensemble de points mesurés**
  - Le critère conduisant à cet optimum est la **minimisation des distances quadratiques entre les points mesurés et la droite optimum**.
  - Il s'agit donc d'un problème **d'optimisation**

Régression linéaire

- **Principe**





Optimisation de la fonctionnelle J(a,b) représentant l'ensemble des droites possibles

Régression linéaire

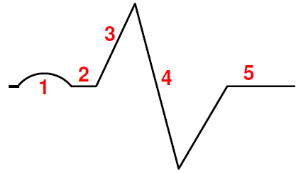
- **Principe**
  - **Autrement dit :**
    - On observe *N* individus (*i* = 1, ..., *N*).
    - Et on suppose que pour tout *i* :

$$Y_i = ax_i + b + E_i$$

    - Et on cherche à estimer a,b et la variance associée aux Y_i (qui est aussi la variance de la variable aléatoire *E*)

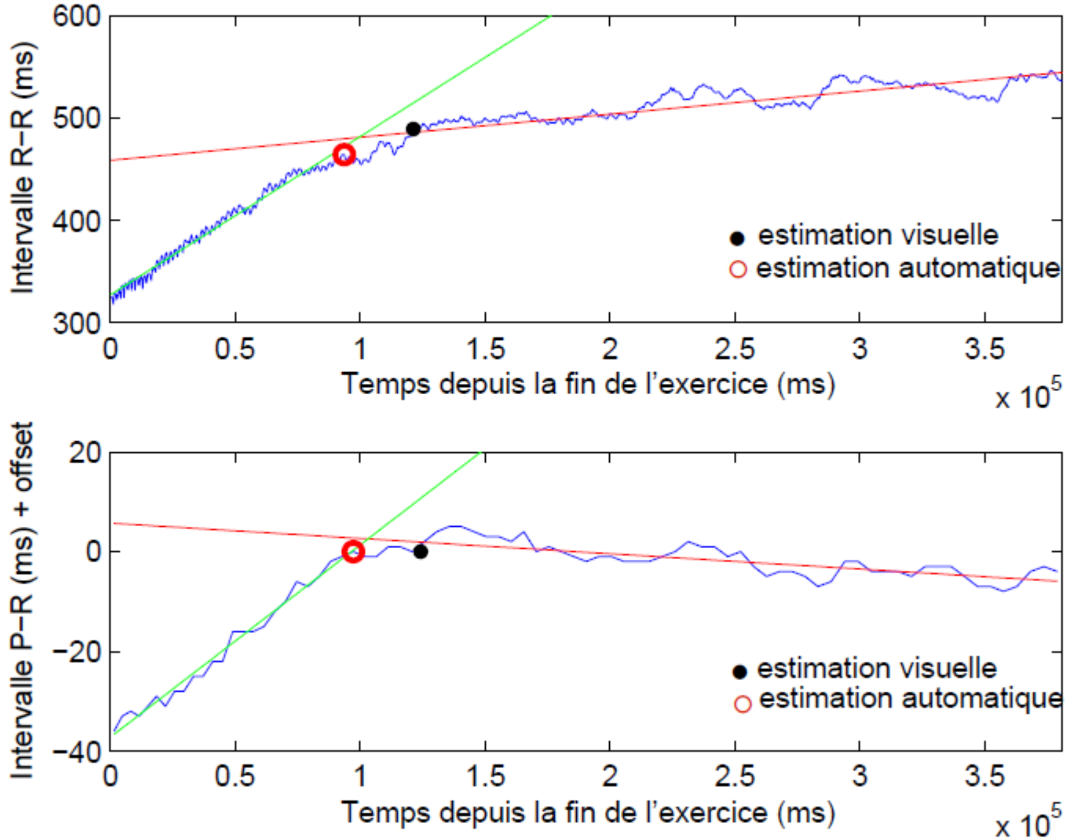
Régression linéaire

- **Remarques**
  - Le principe de régression linéaire s’applique à des grandeurs statiques (qui n’évoluent pas dans le temps)
  - Le filtre de Wiener est une évolution de ce principe mais à des phénomènes évoluant dans le temps
  - Du point de vue optimisation il s’agit d’un problème quadratique convexe, il y a donc nécessairement un minimum et il est unique !



Régression linéaire

- **Exemple sur l'estimation des intervalles ECG**



Régression linéaire

- **2 approches possibles pour l'estimation**
  - *1. Approche des moindres carrés*
    - On cherche les valeurs de *a* et *b* qui minimisent **la somme des carrés des résidus**, *i.e.* les écarts entre les observations (Y_i) et les prédictions ($ax_i + b$) du modèle.

$$MSE(a,b) = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \left(ax_i + b\right)\right)^2$$

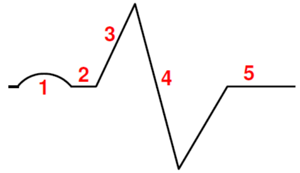
Régression linéaire

- **2 approches possibles pour l'estimation**
  - *2. Approche du Maximum de Vraisemblance (MLE)*
    - Ce critère utilise l'hypothèse de distribution gaussienne des erreurs E_i.

$$MLE(y_i;a,b,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{\left(y_i-(ax_i+b)\right)^2}{2\sigma^2}\right]$$

On retrouve le critère des moindres carrés

Régression linéaire



- **2 approches possibles pour l'estimation**
  - Dans les deux cas, pour l'échantillon *i* pris parmi les *N* disponibles :
    - $E_i = y_i - (ax_i + b)$
    - *La moyenne de E tend vers zéro si le modèle n'est pas biaisé*
    - *La variance de E doit diminuer lorsque les valeurs estimées de a et de b se rapprochent des valeurs exactes*

Régression linéaire

- **2 approches possibles pour l'estimation**
  - <u>Remarques</u>
    - *Les estimateurs de a, b et σ^2 sont des variables aléatoires*
    - *L'approche MLE par l'hypothèse de gaussiannité sur E, permet de caractériser les lois des estimateurs précédents*
    - *On peut alors comparer les estimateurs, définir des intervalles de confiance, etc.*

Régression linéaire

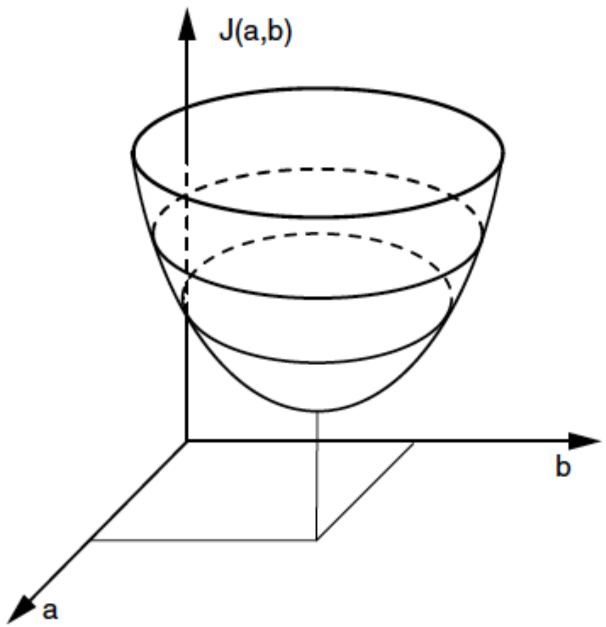
- **Exercice 1**
  - *On se place dans le cadre de l'approche RMS*
  - *Déterminer l'expression analytique de a et de b permettant de minimiser le problème d'estimation considéré entre la variable aléatoire Y et la variable x*
  - *Point de départ : on minimise la fonctionnelle suivante :*

$$MSE(a,b) = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \left(ax_i + b\right)\right)^2$$

Régression linéaire

- **Exercice 1**
  - *Déterminer l'expression analytique de a et de b permettant de minimiser le problème d'estimation considéré entre la variable aléatoire Y et la variable x*

$$MSE(a,b) = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \left(ax_i + b\right)\right)^2$$



Régression linéaire

- **Exercice 1**
  - *Résolution*

On minimise

$$J(a,b) = \frac{1}{N}\sum_{n=0}^{N-1} e^2(n) = \frac{1}{N}\sum_{n=0}^{N-1}\left(y(n) - (ax(n)+b)\right)^2$$

Lorsque l'écart quadratique moyen est minimum

$$\frac{\partial J(a,b)}{\partial a} = 0 \qquad \frac{\partial J(a,b)}{\partial b} = 0$$

Régression linéaire

- **Exercice 1**
  - *Résolution*

Or

$$\begin{aligned}\frac{\partial J(a,b)}{\partial a} &= \frac{1}{N}\sum_{n=0}^{N-1} 2\left(y(n)-(ax(n)+b)\right)\left(0-x(n)-0\right)\\ &= \frac{2}{N}\sum_{n=0}^{N-1}\left(-x(n)y(n)+ax^2(n)+bx(n)\right)\\ &= \frac{2}{N}\left(-\sum_{n=0}^{N-1}x(n)y(n)+a\sum_{n=0}^{N-1}x^2(n)+b\sum_{n=0}^{N-1}x(n)\right)\end{aligned}$$

Régression linéaire

- **Exercice 1**
  - *Résolution*

Et

$$
\begin{aligned}
\frac{\partial J(a,b)}{\partial b} &= \frac{1}{N}\sum_{n=0}^{N-1} 2\left(y(n)-(ax(n)+b)\right)(0-0-1) \\
&= \frac{2}{N}\sum_{n=0}^{N-1}\left(-y(n)+ax(n)+b\right) \\
&= \frac{2}{N}\left(-\sum_{n=0}^{N-1} y(n)+a\sum_{n=0}^{N-1} x(n)+\sum_{n=0}^{N-1} b\right)
\end{aligned}
$$

Régression linéaire

- **Exercice 1**
  - *Résolution*

On obtient donc deux équations :

$$a\frac{1}{N}\sum_{n=0}^{N-1}x^2(n)+b\frac{1}{N}\sum_{n=0}^{N-1}x(n)=\frac{1}{N}\sum_{n=0}^{N-1}x(n)y(n)$$

$$a\frac{1}{N}\sum_{n=0}^{N-1}x(n)+\frac{1}{N}\sum_{n=0}^{N-1}b=\frac{1}{N}\sum_{n=0}^{N-1}y(n)$$

Régression linéaire

- **Exercice 1**
  - *Résolution*

En plus compact

$$a\mu_x + b = \mu_y$$

$$a\mu_{x^2} + b\mu_x = \mu_{xy}$$

Ou encore

$$\begin{pmatrix} \mu_x & 1 \\ \mu_{x^2} & \mu_x \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \mu_y \\ \mu_{xy} \end{pmatrix}$$

Régression linéaire

- **Exercice 1**
  - *Résolution*

Et la solution est donc

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \mu_x & 1 \\ \mu_{x^2} & \mu_x \end{pmatrix}^{-1} \begin{pmatrix} \mu_y \\ \mu_{xy} \end{pmatrix}$$

Analytiquement :

$$a = \frac{\mu_x \mu_y - \mu_{xy}}{\mu_x^2 - \mu_{x^2}}$$

$$b = \frac{\mu_x \mu_{xy} - \mu_y \mu_{x^2}}{\mu_x^2 - \mu_{x^2}}$$

Régression linéaire

- **Exercice 1**
  - *Résolution*

$$a = \frac{\mu_x \mu_y - \mu_{xy}}{\mu_x^2 - \mu_{x^2}}$$

$$b = \frac{\mu_x \mu_{xy} - \mu_y \mu_{x^2}}{\mu_x^2 - \mu_{x^2}}$$

Ou de manière plus classique

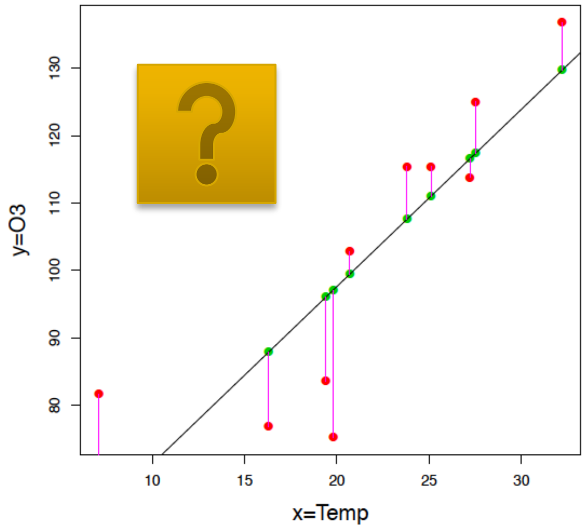
$$\hat{a} = \frac{C_{xy}}{\sigma_x^2} \quad \text{et} \quad \hat{b} = \mu_y - \hat{a}\mu_x$$

Régression linéaire

- **Exercice 2**
  - *Exemple*

Temp.	23.8	16.3	27.2	7.1	25.1	27.5	19.4	19.8	32.2	20.7
O3	115.4	76.8	113.8	81.6	115.4	125.0	83.6	75.2	136.8	102.8

Déterminer les coefficients *a* et *b* associés à ce problème de régression linéaire



Régression linéaire

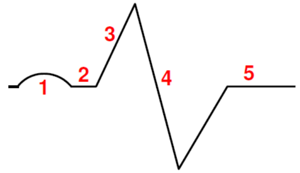
- **Exercice 2**
  - *Résolution*

Temp.	23.8	16.3	27.2	7.1	25.1	27.5	19.4	19.8	32.2	20.7
O3	115.4	76.8	113.8	81.6	115.4	125.0	83.6	75.2	136.8	102.8

On utilise les résultats de l'exercice précédent

$$\hat{a} = \frac{C_{xy}}{\sigma_x^2} \qquad \hat{b} = \mu_y - \hat{a}\mu_x$$

Régression linéaire

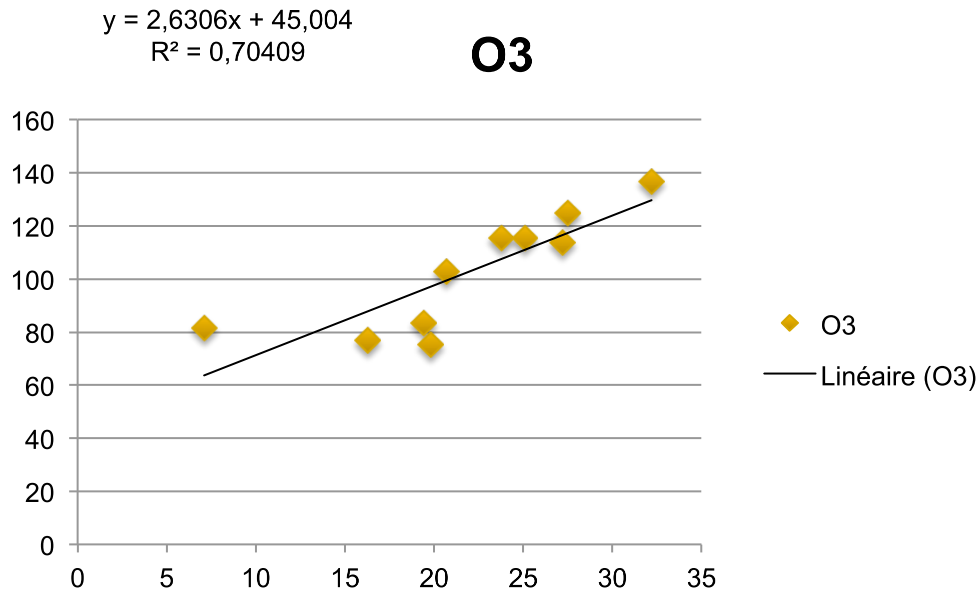


- **Exercice 2**
  - *Résolution*

$$\hat{a} = \frac{C_{xy}}{\sigma_x^2} \qquad \hat{b} = \mu_y - \hat{a}\mu_x$$

Temp.	23.8	16.3	27.2	7.1	25.1	27.5	19.4	19.8	32.2	20.7
O3	115.4	76.8	113.8	81.6	115.4	125.0	83.6	75.2	136.8	102.8

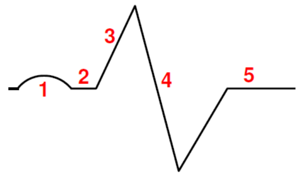
T	23,8
O3	115,4
T*T	566,44
T*O3	2746,52
moy_T	21,91
moy_O3	102,64
moy_T_O3	2365,084
moy_T_T	524,237
CovT_O3	129,1573333
std_T	49,09877778
a	2,630561069
b	45,00440699



Plan

- Introduction
- Rappels sur l'estimation statistique
- Régression linéaire
- Filtrage de Wiener

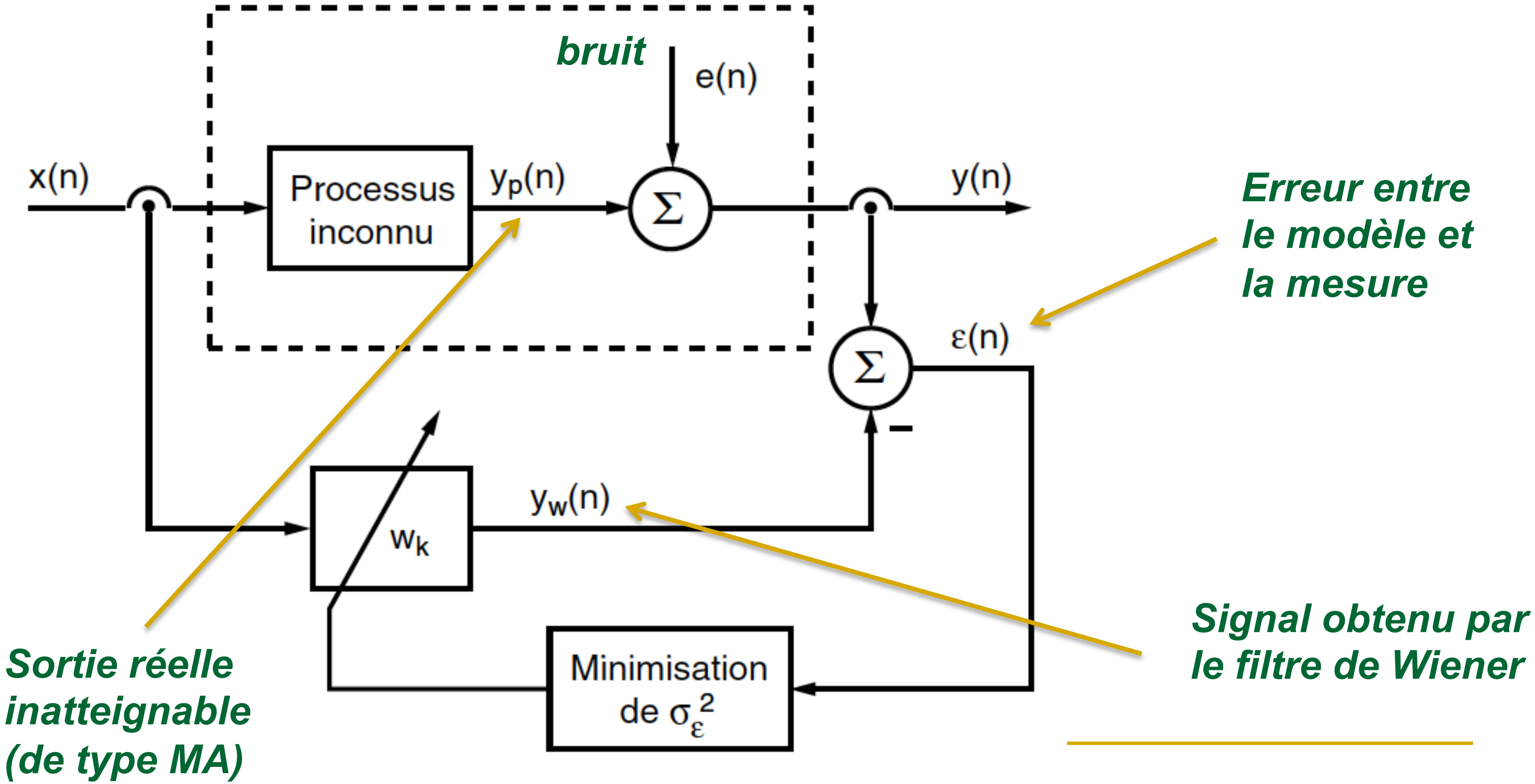
Filtrage de Wiener



- **Contexte**
  - On cherche à supprimer l'influence d'un bruit dont le spectre se superpose à celui du signal d'intérêt
  - Le filtre de Wiener utilise la propriété de stationnarité des signaux aléatoires
  - **Hypothèse :** le signal recherché peut se modéliser à l'aide d'un modèle MA *(Moving Average)*

Filtrage de Wiener

- **Illustration et notations**



Filtrage de Wiener

- **Rappels : Modèle MA, AR**
  - La nature du modèle du signal de sortie dépend du type de la fonction de transfert en z associé au système considéré
  - Plus précisément les modèles MA (moyenne glissante) correspondent aux sorties des filtres de type FIR

$$H(Z) = \frac{Y(Z)}{X(Z)} = \sum_{i=0}^{N-1} b_i . Z^{-i}$$

Filtrage de Wiener

- **Rappels : Modèle MA, AR**
  - La nature du modèle du signal de sortie dépend du type de la fonction de transfert en z associé au système considéré
  - Plus précisément les modèles AR (autorégressif) correspondent aux sorties des filtres de type RII tel que

$$H(Z) = \frac{Y(Z)}{X(Z)} = \frac{1}{\sum_{k=0}^{M-1} a_i . Z^{-k}}$$

Filtrage de Wiener

- **Rappels : Modèle MA, AR**
  - La nature du modèle du signal de sortie dépend du type de la fonction de transfert en z associé au système considéré

  - On parle de modèle ARMA lorsque

$$H(Z) = \frac{Y(Z)}{X(Z)} = \frac{\sum_{i=0}^{N-1} b_i . Z^{-i}}{\sum_{k=0}^{M-1} a_i . Z^{-k}}$$

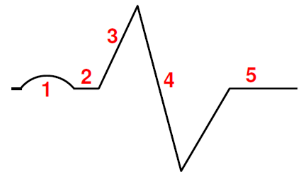
Filtrage de Wiener

- **Problématique**
  - Le signal $y_p(n)$ est de type MA d'ordre p

$$y_p(n) = \sum_{k=0}^{p-1} w_k . x(n-k)$$

  - Le but poursuivi est de trouver les coefficients w_k du modèle MA à partir de la mesure des signaux d'entrée $x(n)$ et de sortie $y(n)$.

Filtrage de Wiener



- **Problématique**
  - La recherche d'une solution consiste à rendre $y_w(n)$ aussi proche que possible du signal $y_p(n)$…
  - … en minimisant **l'erreur quadratique moyenne** par ajustement des coefficients w_k.
  - On revient donc à un problème proche de la régression linéaire.

Filtrage de Wiener

- **Remarque importante**

Si la solution exacte est trouvée, le signal d'écart *ε(n)* n'est pas nul, **mais égal au bruit de la mesure *e(n)***.

Filtrage de Wiener

- **Résolution**
  - *Pour simplifier les calculs, on considèrera un processus MA d'ordre 3*
  - *On notera* ***W*** *le vecteur associé aux 3 paramètres recherchés*

$$W = \begin{pmatrix} w_0 \\ w_1 \\ w_2 \end{pmatrix}$$

Filtrage de Wiener

- **Résolution**
  - *Le signal estimé $y_w(n)$ peut alors s'écrire*

$$y_w(n) = w_0 x(n) + w_1 x(n-1) + w_2 x(n-2)$$

  - *On cherche à minimiser la fonctionnelle suivante*

$$J = \frac{1}{N} \sum_{n=0}^{N-1} \left(y(n) - y_w(n) \right)^2$$

Filtrage de Wiener

- **Exercice ☺**
  - **Montrer que la solution de ce problème d'optimisation est en se limitant à 3 coefficients :**

$$W = R_{XX}^{-1} r_{xy}$$

Avec R_{XX} la matrice d'autocorrélation associée à la variable *x et*

$$r_{xy}(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)\,y(n+k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n-k)\,y(n) = r_{yx}(-k)$$

la corrélation associée aux variables *x et y*

Filtrage de Wiener

- **Résolution**

  - De la même manière que pour la régression linéaire, J est minimale lorsque les 3 dérivées par rapport aux w_i s'annulent

  - Soit

$$
\begin{aligned}
\frac{\partial J}{\partial w_0} &= \frac{2}{N}\sum_{n=0}^{N-1}\left(y(n) - w_0 x(n) - w_1 x(n-1) - w_2 x(n-2)\right)\left(-x(n)\right) \\
&= -\frac{2}{N}\sum_{n=0}^{N-1}\left(x(n)y(n) - w_0 x(n)x(n) - w_1 x(n)x(n-1) - w_2 x(n)x(n-2)\right)
\end{aligned}
$$

Filtrage de Wiener

- **Résolution**

  - 3 équations :

$$
\begin{aligned}
\frac{\partial J}{\partial w_0} &= \frac{2}{N}\sum_{n=0}^{N-1}\left(y(n) - w_0 x(n) - w_1 x(n-1) - w_2 x(n-2)\right)\left(-x(n)\right) \\
&= -\frac{2}{N}\sum_{n=0}^{N-1}\left(x(n)y(n) - w_0 x(n)x(n) - w_1 x(n)x(n-1) - w_2 x(n)x(n-2)\right)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial w_1} &= \frac{2}{N}\sum_{n=0}^{N-1}\left(y(n) - w_0 x(n) - w_1 x(n-1) - w_2 x(n-2)\right)\left(-x(n-1)\right) \\
&= -\frac{2}{N}\sum_{n=0}^{N-1}\left(x(n-1)y(n) - w_0 x(n-1)x(n) - w_1 x(n-1)x(n-1) - w_2 x(n-1)x(n-2)\right)
\end{aligned}
$$

Filtrage de Wiener

- **Résolution**
  - et

$$
\begin{aligned}
\frac{\partial J}{\partial w_2} &= \frac{2}{N}\sum_{n=0}^{N-1}\left(y(n)-w_0x(n)-w_1x(n-1)-w_2x(n-2)\right)\left(-x(n-2)\right)\\
&= -\frac{2}{N}\sum_{n=0}^{N-1}\left(x(n-2)y(n)-w_0x(n-2)x(n)-w_1x(n-2)x(n-1)-w_2x(n-2)x(n-2)\right)
\end{aligned}
$$

Filtrage de Wiener

- **Résolution**
  - En utilisant l'intercorrélation, il vient que

$$
\begin{aligned}
\frac{\partial J}{\partial w_0} &= -2\left(r_{xy}(0) - w_0 r_{xx}(0) - w_1 r_{xx}(-1) - w_2 r_{xx}(-2)\right) \\
\frac{\partial J}{\partial w_1} &= -2\left(r_{xy}(+1) - w_0 r_{xx}(+1) - w_1 r_{xx}(0) - w_2 r_{xx}(-1)\right) \\
\frac{\partial J}{\partial w_2} &= -2\left(r_{xy}(+2) - w_0 r_{xx}(+2) - w_1 r_{xx}(+1) - w_2 r_{xx}(0)\right)
\end{aligned}
$$

Filtrage de Wiener

- **Résolution**

  - On obtient donc un système de 3 équations à 3 inconnues

$$w_0 r_{xx}(0) + w_1 r_{xx}(-1) + w_2 r_{xx}(-2) = r_{xy}(0)$$

$$w_0 r_{xx}(+1) + w_1 r_{xx}(0) + w_2 r_{xx}(-1) = r_{xy}(+1)$$

$$w_0 r_{xx}(+2) + w_1 r_{xx}(+1) + w_2 r_{xx}(0) = r_{xy}(+2)$$

Filtrage de Wiener

- **Résolution**
  - Soit sous forme matricelle

$$\begin{pmatrix} r_{xx}(0) & r_{xx}(-1) & r_{xx}(-2) \\ r_{xx}(+1) & r_{xx}(0) & r_{xx}(-1) \\ r_{xx}(+2) & r_{xx}(+1) & r_{xx}(0) \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} r_{xy}(0) \\ r_{xy}(1) \\ r_{xy}(2) \end{pmatrix}$$

Matrice d'autocorrélation

Et donc

$$R_{XX} W = r_{xy}$$

Filtrage de Wiener

- **Résolution**
  - Et finalement (CQFD)

$$W = R_{XX}^{-1} r_{xy}$$

Il s’agit de l’équation de Wiener-Hopf

Filtrage de Wiener

- **Généralisation**
  - Ce résultat se généralise à tout ordre *p* associé au vecteur *W*

$$W = \begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_{p-1} \end{pmatrix}$$

auquel on associe le vecteur X(n) tel que

$$X(n) = \begin{pmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-p+1) \end{pmatrix}$$

Filtrage de Wiener

- **Généralisation**

  - On obtient alors

$$y_w(n) = \sum_{i=0}^{p-1} w_i . x(n-i) = W^T X(n) = X(n)^T W$$

  - L'erreur quadratique est donnée par

$$\varepsilon(n)^2 = \left(y(n) - X(n)^T W \right)^2$$

Filtrage de Wiener

- **Généralisation**
  - On minimise alors l'espérance de l'erreur quadratique moyenne

$$J(W) = E\left[\varepsilon(n)^2\right] = E\left[\left(y(n) - X(n)^T W\right)^2\right]$$

  - Ce que l'on peut encore écrire

$$J(W) = E\left[y(n)^2\right] - 2E\left[y(n)X(n)^T W\right] + E\left[W^T X(n)X(n)^T W\right]$$

Filtrage de Wiener

- **Généralisation**
  - Ce que l'on peut écrire

$$J(W)=E\left[y(n)^2\right]-2E\left[y(n)X(n)^T W\right]+E\left[W^T X(n)X(n)^T W\right]$$

  - Ou encore

$$J(W)=r_{yy}(0)-2r_{xy}^T W+W^T R_{XX}W$$

Filtrage de Wiener

- **Généralisation** $J(W) = r_{yy}(0) - 2r_{xy}^{T}W + W^{T}R_{XX}W$

  - On minimise donc *J* dans la direction *W, ce qui revient à annuler le gradient de J dans cette direction.*

  - On résout donc

$$\frac{\partial J(W)}{\partial W} = -2r_{xy} + 2R_{XX}W = 0$$

  - Soit

$$W = R_{XX}^{-1}r_{xy}$$

Filtrage de Wiener

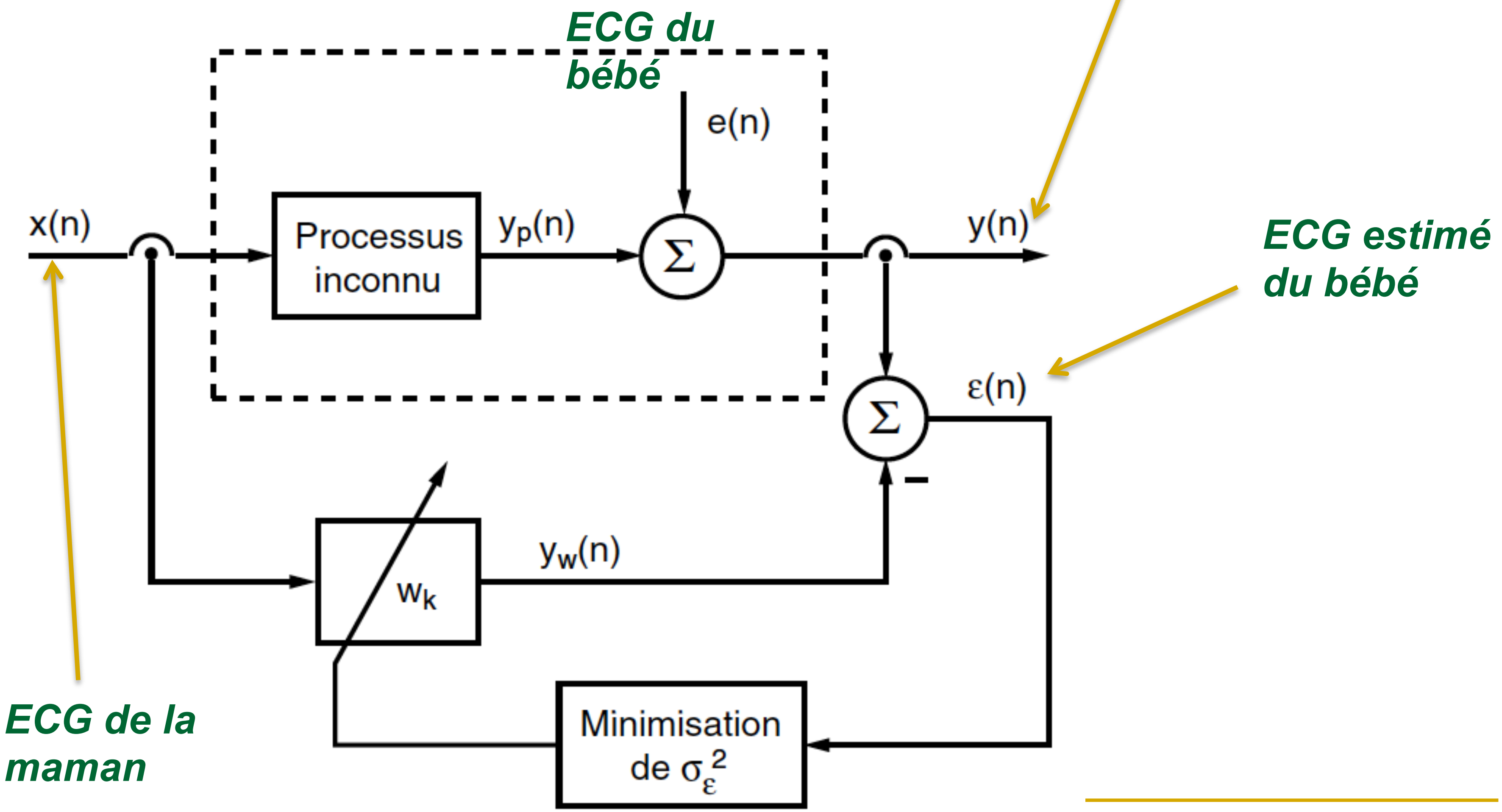
- **Quelques applications**
  - Identification de processus (on cherche les ***w*** qui modélisent au mieux le processus inconnu)
  - La prédiction linéaire (on estime une valeur à venir en prenant en compte les valeurs précédentes)
  - **La suppression d'un signal perturbateur**

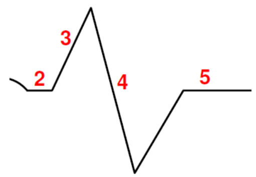
Filtrage de Wiener

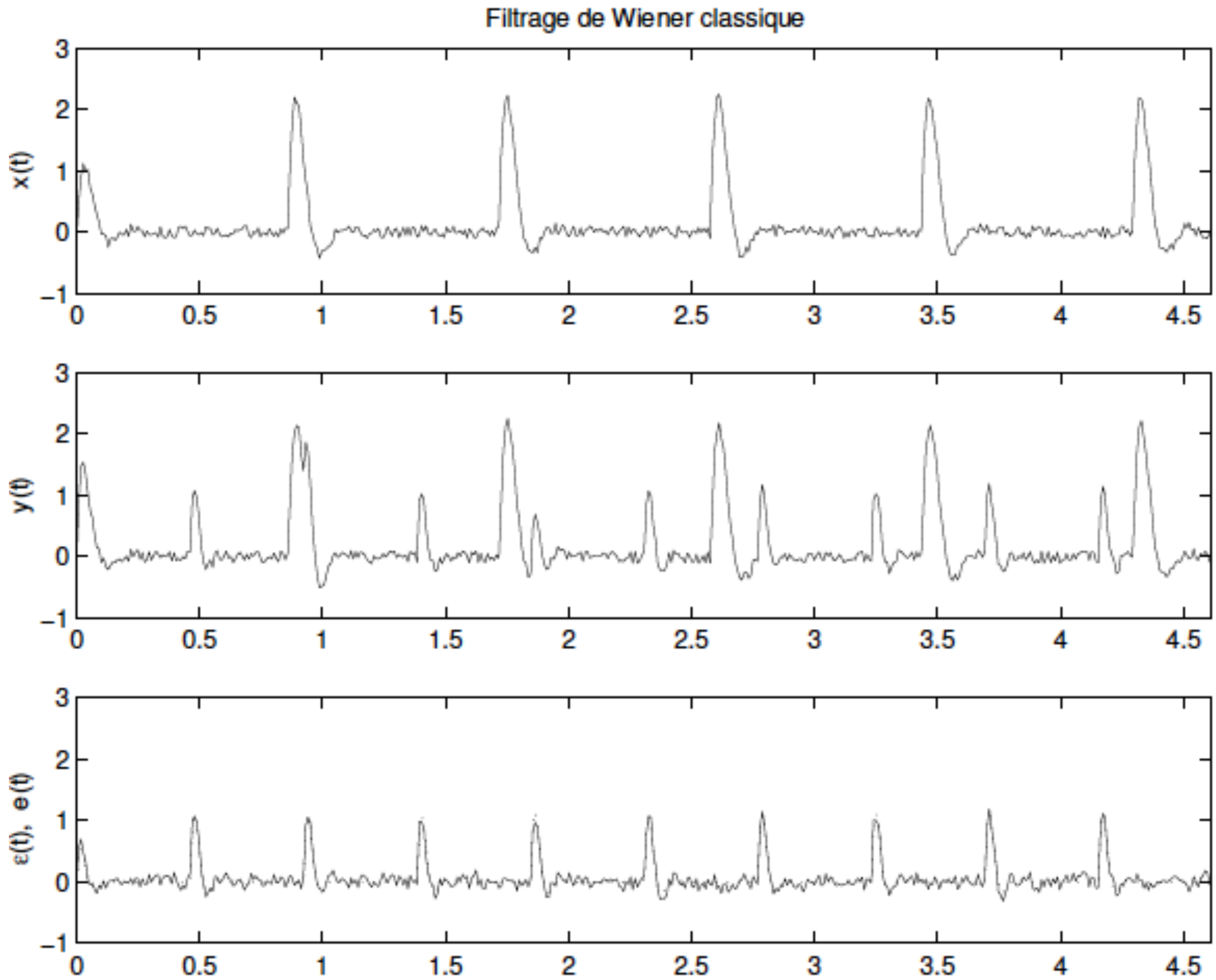
- **Exemple**
  - Extraction de l'ECG d'un fœtus superposé à celui de la maman (mesure au moyen d'une sonde Doppler)
  - Dans ce contexte, le signal d'entrée est l'ECG de la maman (facilement mesurable au moyen d'un système d'acquisition classique)
  - Le signal recherché est donc paradoxalement l'erreur e(n)

Filtrage de Wiener

- **Illustration et notations**







ECG maman

Signaux superposés (Doppler)

Estimation de l'ECG du bébé

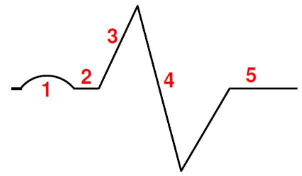
Filtrage de Wiener

- **Limites**
  - Le calcul de la matrice d'autocorrélation et de la corrélation est gourmand en calcul
  - L'inversion de la matrice d'autocorrélation est également gourmande
  - Si l'hypothèse de stationnarité n'est pas respectée (ce qui est assez souvent le cas), la matrice d'autocorrélation et la corrélation évoluent dans le temps ; il faut donc à chaque itération les recalculer
  - **Solution :** le filtrage adaptatif

Plan

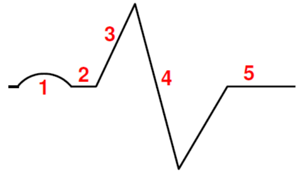
- Introduction
- Rappels sur l’estimation statistique
- Régression linéaire
- Filtrage de Wiener
- Filtrage adaptatif

Filtrage adaptatif



- **Définition**
  - Un filtre adaptatif est un système numérique dont les coefficients se modifient eux mêmes en fonction des signaux extérieurs.
  - Il est utilisé chaque fois qu’un environnement est mal connu ou changeant ou pour supprimer des perturbations situées dans le domaine de fréquences du signal utile…
  - …**ce que les filtres classiques ne peuvent pas faire.**

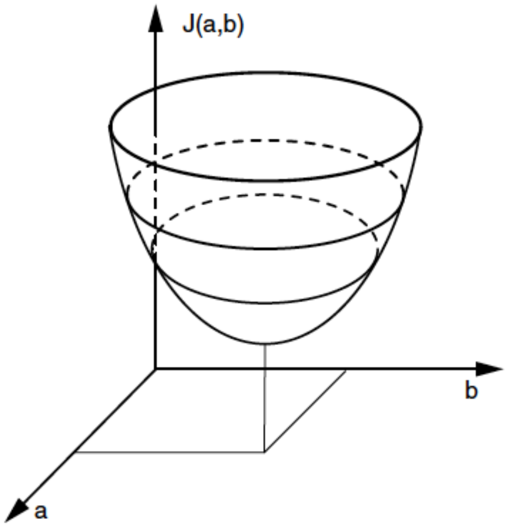
Filtrage adaptatif



- **Principe**
  - Un filtre adaptatif est constitué de deux parties distinctes :
    - un filtre numérique à coefficients ajustables ;
    - un algorithme de modification des coefficients basé sur un critère d'optimisation.
      - ***Exemples :*** Algorithme récursif des moindres carrés, Algorithme récursif normalisé

Filtrage adaptatif

- **Algorithme récursif des moindres carrés (RLMS)**
  - **Principe :** Nous avons vu précédemment que le filtrage de Wiener peut se voir comme un problème d'optimisation avec une fonction « coût » de type paraboloïde



Avec *J* l'erreur quadratique moyenne associée au problème

Filtrage adaptatif

- **Algorithme récursif des moindres carrés (RLMS)**
  - Le minimum s'obtient en descendant le long de la pente de la paraboloïde jusqu'à annulation de la dérivée de la fonction de coût dans la direction *W (les coefficients optimaux du filtre)*

$$\frac{\partial J(W)}{\partial W} = -2r_{xy} + 2R_{XX}W$$

Et lorsque la dérivée s'annule

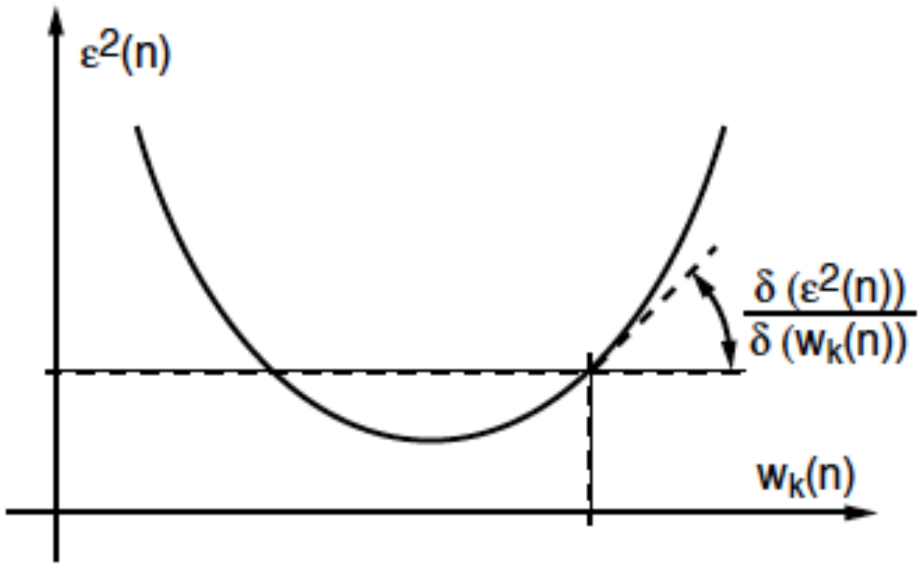
$$W = R_{XX}^{-1} r_{xy}$$

Filtrage adaptatif

- **Algorithme récursif des moindres carrés (RLMS)**
  - De manière heuristique, on imagine bien que cette solution peut être atteinte récursivement…
  - …en corrigeant les valeurs des coefficients $\boldsymbol{w_k}$ en chaque instant $\boldsymbol{n}$ dans le sens opposé à l'évolution de l'erreur quadratique par rapport au vecteur des coefficients $\boldsymbol{W}$

Filtrage adaptatif

- **Algorithme récursif des moindres carrés (RLMS)**
  - ***Illustration graphique :***



Soit

$$W(n) = W(n-1) - \frac{\gamma}{2}\left(\frac{\partial \varepsilon^2(n)}{\partial W}\right)$$

On parle de descente de gradient

Rq : γ est la facteur de pondération du gradient

Filtrage adaptatif

- **Algorithme récursif des moindres carrés (RLMS)**
  - ***Exercice :***
    - *Montrer que l'expression de la fonction de mise à jour des coefficients w peut s'écrire*

$$W(n) = W(n-1) + \gamma\varepsilon(n)X(n)$$

Filtrage adaptatif

- **Algorithme récursif des moindres carrés (RLMS)**
  - ***Exercice (Solution):***

$$\varepsilon^2(n) = \left(y(n) - \sum_{i=0}^{p-1} w_i x(n-i) \right)^2 = \left(y(n) - X(n)^T W \right)^2$$

Et donc

$$\frac{\partial \varepsilon^2(n)}{\partial W} = 2\varepsilon(n) \frac{\partial \varepsilon(n)}{\partial W} = -2\varepsilon(n) X(n)$$

Filtrage adaptatif

- **Algorithme récursif des moindres carrés (RLMS)**
  - ***Exercice (Solution):***

Finalement, en remplaçant

$$W(n) = W(n-1) + \gamma \varepsilon(n) X(n)$$

Il s'agit de l'algorithme *Recursive Least Mean Square (RLMS)*

Filtrage adaptatif

- **Algorithme récursif des moindres carrés (RLMS)**
  - ***Ce dont on a besoin***
    - *Le vecteur des p coefficients du filtre à l'instant n-1*

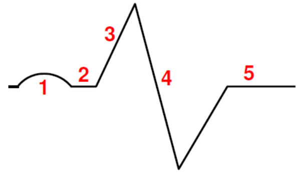
$$W(n-1) = \left[w_0(n-1), w_1(n-1), \cdots, w_{p-1}(n-1)\right]^T$$

Filtrage adaptatif

- **Algorithme récursif des moindres carrés (RLMS)**
  - ***Ce dont on a besoin***
    - *Les p dernières valeurs du signal d'entrée*

$$X(n) = \left[x(n), x(n-1), \cdots, x(n-p+1)\right]^T$$

Filtrage adaptatif



- **Algorithme récursif des moindres carrés (RLMS)**
  - ***Ce dont on a besoin***
    - *La valeur du signal de sortie à l'instant n, y(n)*
    - *Le gain d'adaptation γ de l'algorithme récursif (<<1)*

Filtrage adaptatif

- **Algorithme récursif des moindres carrés (RLMS)**
  - ***Remarque sur le gain***
    - *Son choix est empirique et pas toujours évident*
      - *Trop petit, la vitesse de convergence est très lente*
      - *Trop grand, l'algorithme peut osciller longtemps autour de la valeur optimale…*
      - *… et même entraîner une divergence de l'algorithme*

Filtrage adaptatif

- **Algorithme récursif des moindres carrés (RLMS)**
  - ***Avantages/Inconvénients***
    - **+** *Rapide à mettre en œuvre*
    - **+** *Peu de calcul*

    - **-** *Lente convergence*
    - **-** *Instabilité (oscillation, divergence) si le gain est trop grand*

Nécessité de proposer une solution d'adaptation automatique du gain

Filtrage adaptatif

- **Algorithme récursif normalisé (NLMS)**
  - ***Principe***
    - ***1975 :*** *Widrow et al montre qu'on peut adapter le gain de manière stable entre 0 et 1 en le normalisant :*
      - *Par le nombre p de paramètres du vecteur W*
      - ***Et** par la puissance ou par la variance du signal d'entrée*

Filtrage adaptatif

- **Algorithme récursif normalisé (NLMS)**
  - ***Cas d'un signal stationnaire***
    - *En prenant stricto sensu la normalisation proposée par Widrow et al :*

$$\gamma = \frac{\gamma_0}{p.\sigma_x^2}$$

La valeur initiale est généralement choisie à 0.1, et la valeur s'adapte à la puissance du signal (via la variance)

Filtrage adaptatif

- **Algorithme récursif normalisé (NLMS)**
  - ***Cas d'un signal stationnaire***
    - *Cette solution directe présente l'inconvénient d'avoir un gain qui peut augmenter de manière non maîtrisé*

$$\lim_{\sigma_x \to 0} \gamma = \lim_{\sigma_x \to 0} \frac{\gamma_0}{p.\sigma_x^2} = +\infty$$

Filtrage adaptatif

- **Algorithme récursif normalisé (NLMS)**
  - ***Cas d'un signal stationnaire***
    - *On ajoutera alors un terme constant au dénominateur telle que*

$$\gamma = \frac{\gamma_0}{a + p.\sigma_x^2}$$

a<<1

Filtrage adaptatif

- **Algorithme récursif normalisé (NLMS)**
  - ***Cas d'un signal stationnaire***
    - *L'algorithme s'écrit alors*

$$W(n) = W(n-1) + \frac{\gamma_0}{a + p.\sigma_x^2}\varepsilon(n)X(n)$$

Normalisation, d'où le nom
NLMS (Normalized LMS)

Filtrage adaptatif

- **Algorithme récursif normalisé (NLMS)**
  - ***Cas d'un signal non stationnaire***
    - *Dans ce cas, il n'est plus possible d'estimer la puissance du signal sans prendre en compte le temps*
    - *Il faut donc l'estimer en tout instant*

$$P_x(n) = \frac{1}{n+1} \sum_{k=0}^{n} x^2(k)$$

Filtrage adaptatif

- **Algorithme récursif normalisé (NLMS)**
  - ***Cas d'un signal non stationnaire***
    - *En pratique on utilisera un filtre passe bas d'ordre 1 à mémoire « courte »*
    - *Exemple de fonction de transfert*

$$H(z) = \frac{1-\lambda}{1-\lambda.z^{-1}} \qquad 0<\lambda<1$$

Filtrage adaptatif

- **Algorithme récursif normalisé (NLMS)**
  - ***Cas d'un signal non stationnaire***
    - *En pratique on utilisera un filtre passe bas d'ordre 1 à mémoire « courte »*
    - *Equation de récurrence associée*

$$s(n) = \left(1 - \lambda\right)e(n) + \lambda s(n-1) \qquad 0<\lambda<1$$

Filtrage adaptatif

- **Algorithme récursif normalisé (NLMS)**
  - ***Cas d'un signal non stationnaire***
    - *En pratique on utilisera un filtre passe bas d'ordre 1 à mémoire « courte »*
    - *Calcul de la puissance de manière récursive*

$$P_x(n) = (1-\lambda)\,x^2(n) + \lambda P_x(n-1) \qquad 0<\lambda<1$$

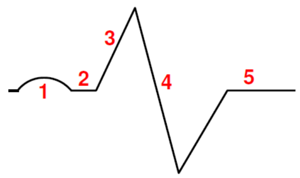
Filtrage adaptatif

- **Algorithme récursif normalisé (NLMS)**
  - ***Cas d'un signal non stationnaire***
    - *En pratique on utilisera un filtre passe bas d'ordre 1 à mémoire « courte »*
    - *Le paramètre λ* fixe l'horizon de mémoire N souhaité

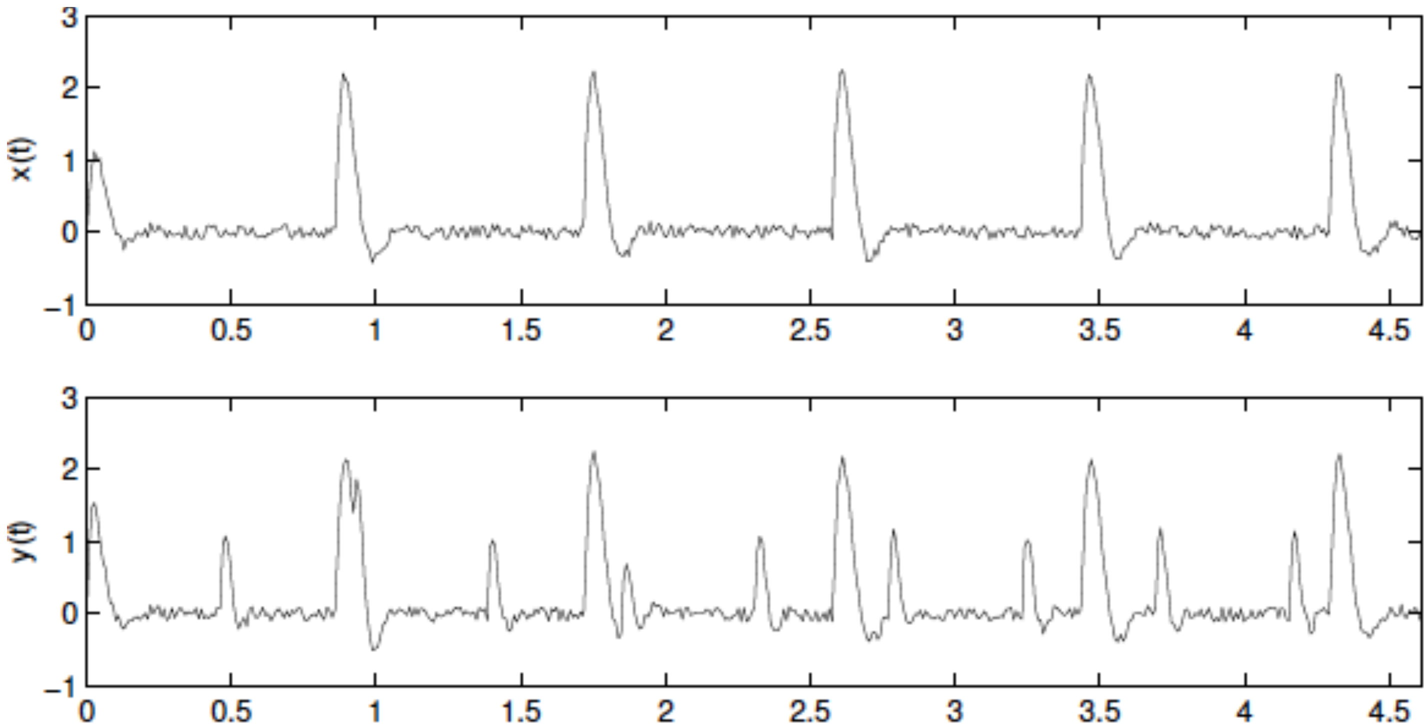
$$N = \frac{3}{|Ln\lambda|}$$

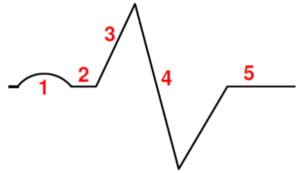
Exemple : *N=30 pour λ=0,9 ; N=150 pour λ=0,98*

Filtrage adaptatif



- **Résultats**
  - ***ECG maman/fœtus : données***

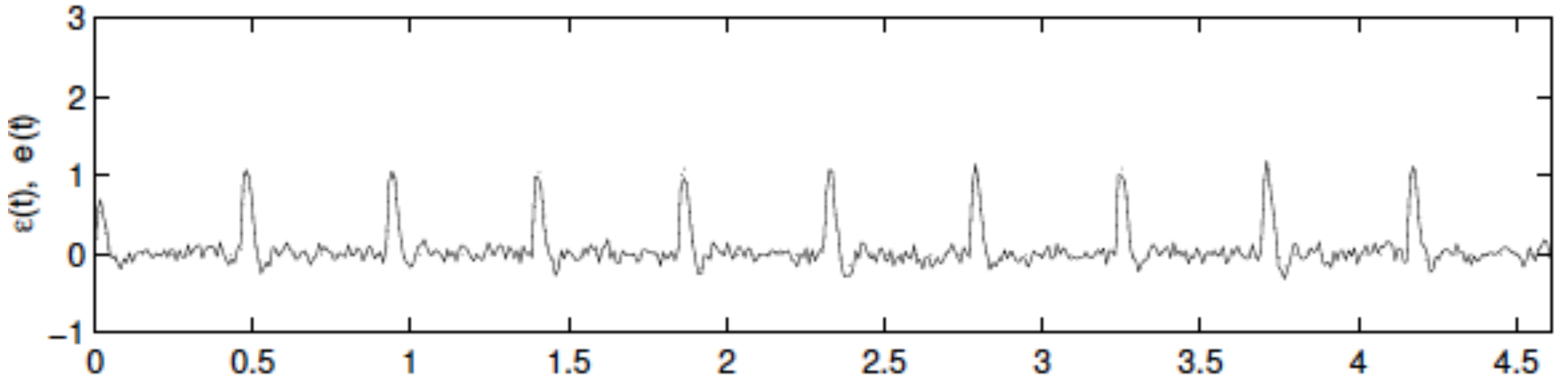




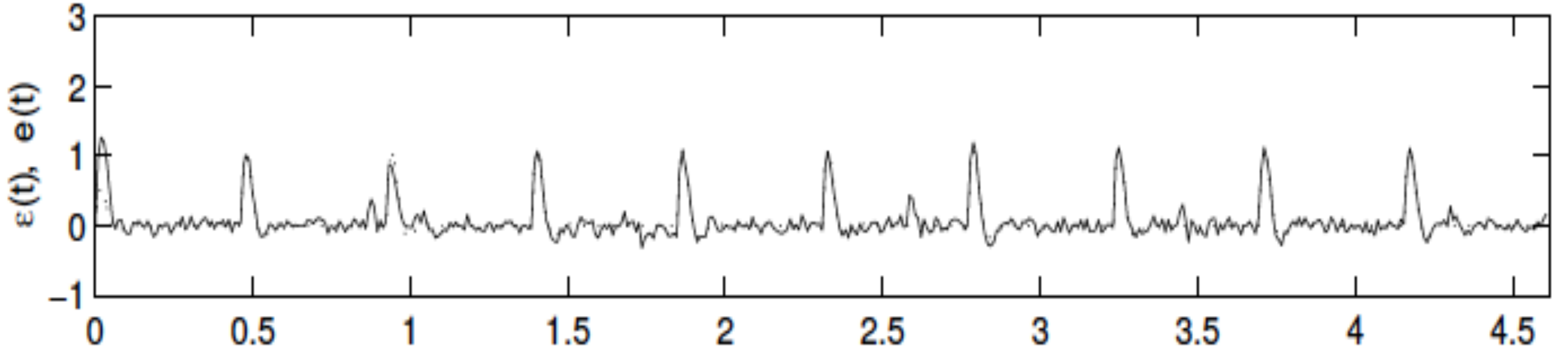
Filtrage adaptatif

- **Résultats**
  - ***ECG maman/fœtus :***

Wiener



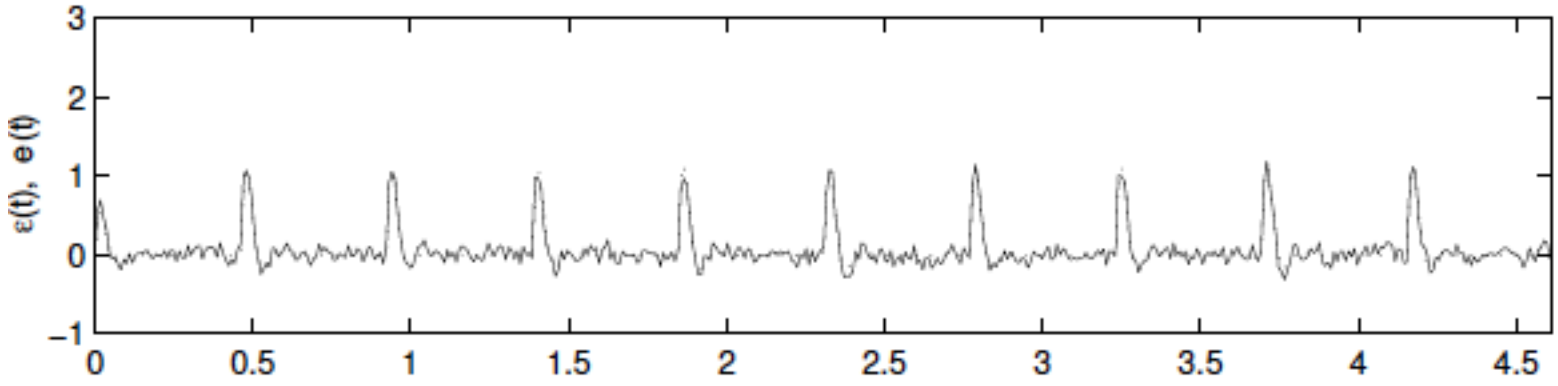
RLMS, γ=0.1



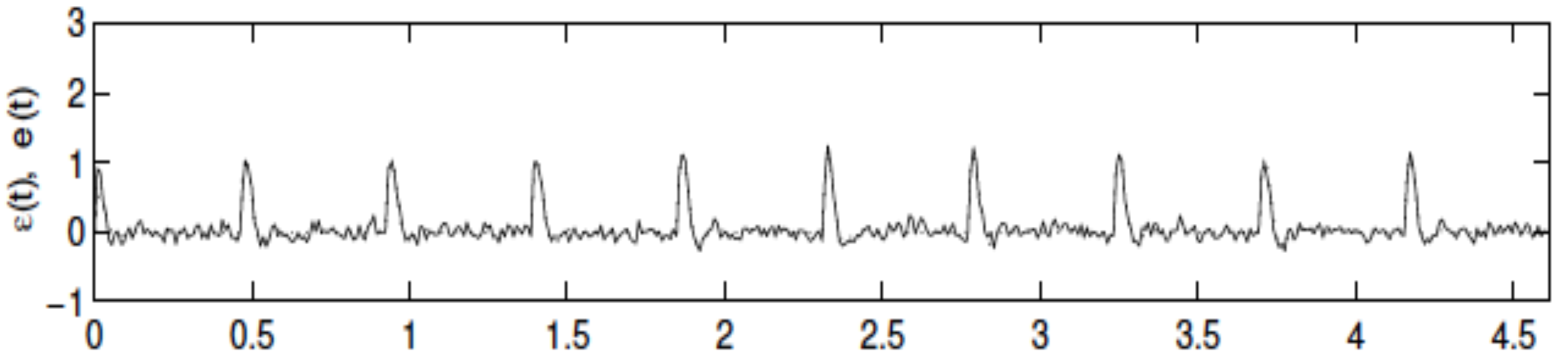
Filtrage adaptatif

- **Résultats**
  - ***ECG maman/fœtus :***

Wiener



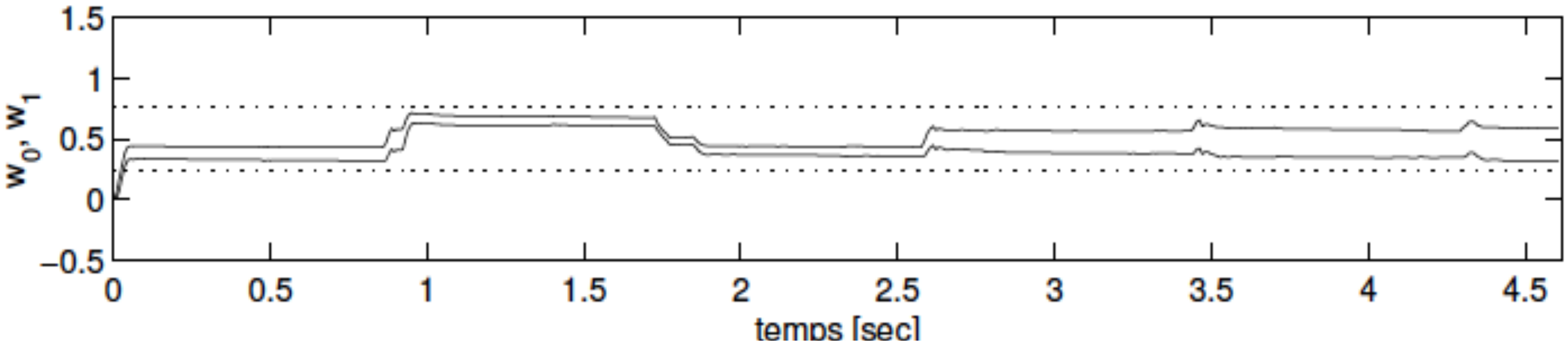
NLMS,
γ_0=0.05,
λ=0,95



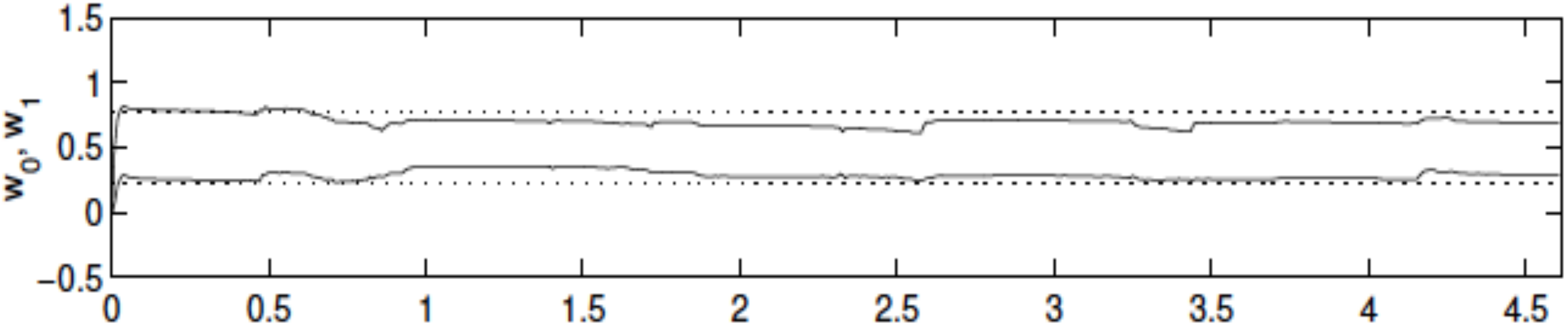
Filtrage adaptatif

- **Résultats**
  - ***ECG maman/fœtus : comparaison RLMS, NLMS***

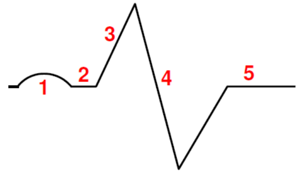
RLMS,
γ=0.1



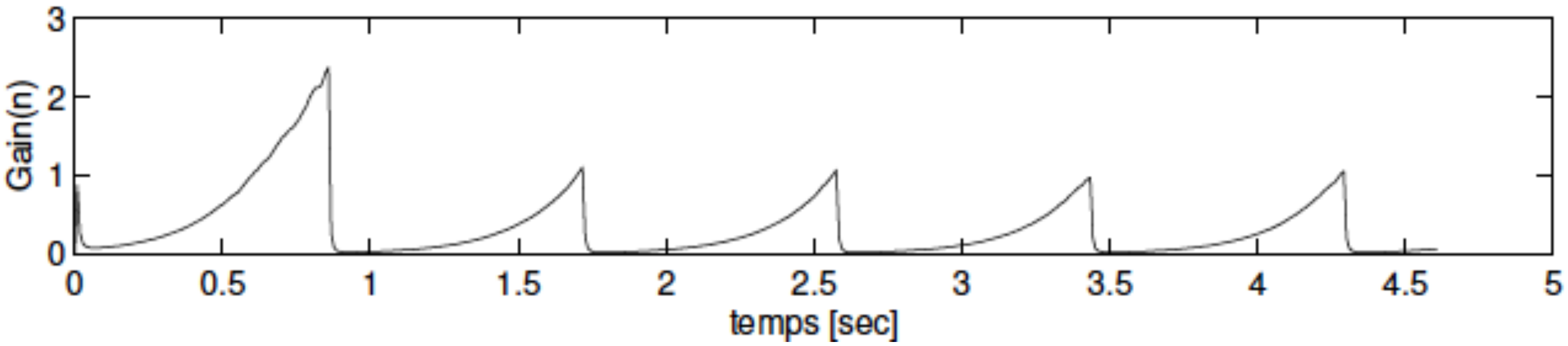
NLMS,
γ_0=0.05
λ=0,95



Filtrage adaptatif

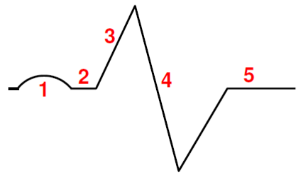


- **Résultats**
  - ***ECG maman/fœtus : comparaison RLMS, NLMS***



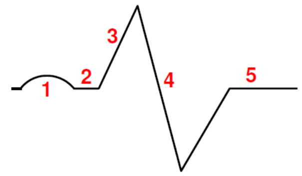
Evolution du gain pour l'algo NLMS, γ_0=0.05, λ=0,95

Filtrage adaptatif



- **Résultats**
  - ***ECG maman/fœtus : comparaison RLMS, NLMS***
    - ***L 'algo NLMS est plus rapide en termes de convergence***
    - ***et plus stable également***
    - ***Le résultat obtenu est très proche de celui obtenu au moyen du filtre optimal de Wiener***

Plan

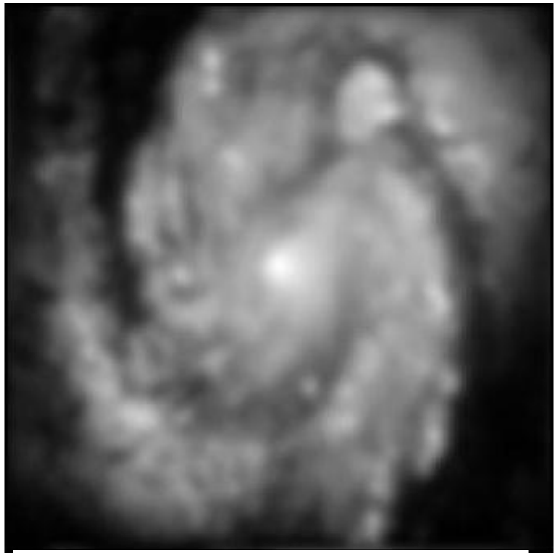


- Introduction
- Rappels sur l'estimation statistique
- Régression linéaire
- Filtrage de Wiener
- Filtrage adaptatif
- Exercice

Exercices

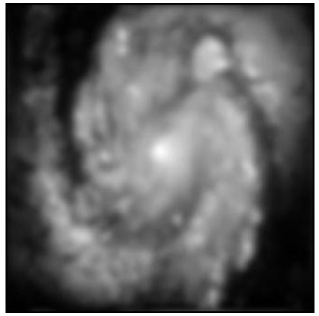
- **Contexte : Restauration d’images bruitées**
  - On considère ici la problématique de restauration d’images astronomiques acquises pas un télescope type Hubble

Image bruitée



Galaxie spirale M100

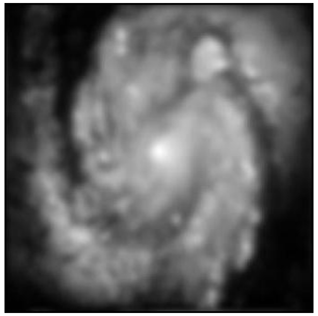
Exercices



- **Problématique**
  - L'image reçue par le CCD est une collection de pixels que l'on rassemble sous la forme d'une matrice ***y***
  - Cette matrice résulte de l'image initiale, ***x***, qui a été convoluée par la réponse impulsionnelle, ***h***, du télescope auquel s'ajoute un vecteur bruit noté ***n (blanc)***

$$y(i,j) = h(i,j) * x(i,j) + n(i,j)$$

Exercices

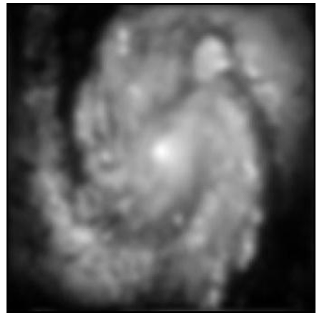


- **Espace des fréquences**
  - ***Question 1 :*** Que devient cette relation dans l'espace de Fourier ?

$$Y(k,l) = H(k,l)X + N(k,l)$$

Rq : Dans cet espace, la matrice *H* est diagonale de valeur propre $\boldsymbol{\lambda_k}$

Exercices

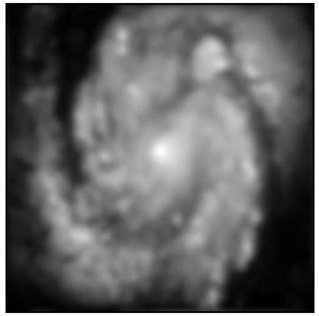


- **Espace des fréquences**
  - Le spectre de puissance de l'image suit souvent une loi de puissance,

$$\left\langle X^2 \right\rangle = C_0 k^{-\alpha}$$

  - Le bruit est souvent un bruit blanc c'est à dire qu'il à la même intensité quel que soit la fréquence spatiale,

$$\left\langle N^2 \right\rangle = C$$

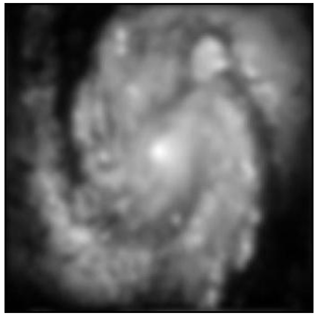


Exercices

- **Espace des fréquences**

  - ***Question 2 :*** Montrer que l'inversion simple de la relation précédente (qui consiste à appliquer la matrice $\boldsymbol{H}^{-1}$ sur les données $\boldsymbol{Y}$ afin de retrouver $\boldsymbol{X}$) conduit, au delà d'une certaine fréquence, à une amplification du bruit.

Exercices



- **Espace des fréquences**
  - ***Solution :*** En inversant simplement la relation précédente, on obtient

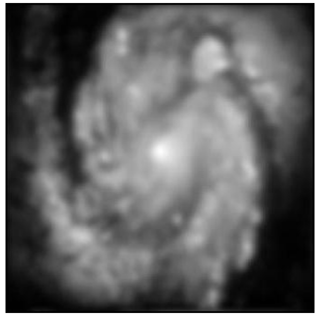
$$H^{-1}Y = H^{-1}HX + H^{-1}N = X + \frac{N}{\lambda_k}$$

  - Comme le spectre de puissance de l'image décroît quand ***k*** augmente et que celui du bruit reste constant, il existe une fréquence limite telle que

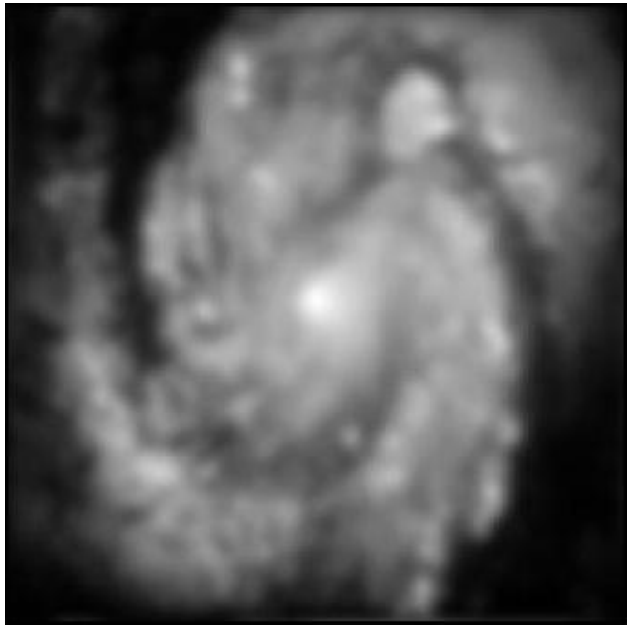
$$\frac{N}{\lambda_k} >> X$$

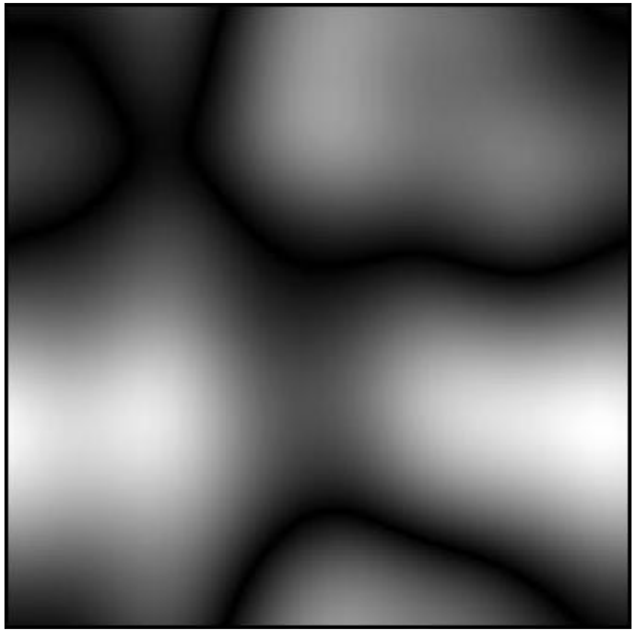
Et donc le bruit est amplifié

Exercices

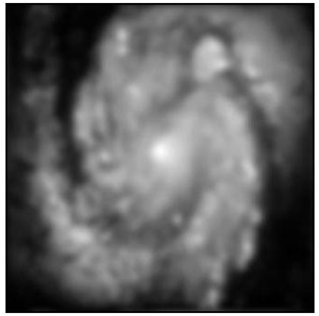


- **Espace des fréquences**
  - ***Restauration par simple inversion***





Exercices



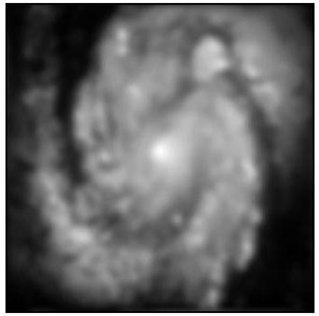
- **Filtre optimal**

  - On cherche donc maintenant à déconvoluer l'image mais aussi à filtrer le bruit.

  - Pour cela, on va chercher le filtre ***R*** à appliquer sur les données ***Y*** qui va minimiser l'écart quadratique moyen entre la vraie image ***X*** et l'image filtrée définie par

$$\hat{X} = RY$$

Exercices



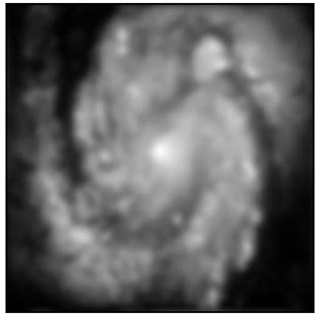
- **Filtre optimal**

  - ***Question 3 :*** En postulant que le bruit et le signal sont décorrélés et que le bruit est iid, montrer que

$$R = \left(H^T C_N^{-1} H + C_X^{-1} \right)^{-1} H^T C_N^{-1}$$

avec $C_X=<X^TX>$ et $C_N=<N^TN>$ les matrices de covariances du signal X et du bruit N

Exercices



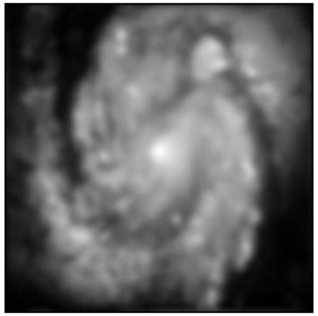
- **Filtre optimal**
  - ***Solution :***

$$S = \left\langle \left(\tilde{X}-X\right)^2 \right\rangle = \left\langle \left(\tilde{X}-X\right)^T \left(\tilde{X}-X\right) \right\rangle = \left\langle (RY-X)^T (RY-X) \right\rangle = \left\langle (\mathrm{R}(AX+N)-X)^T (\mathrm{R}(AX+N)-X) \right\rangle$$

$$\frac{dS}{dR} = 0$$

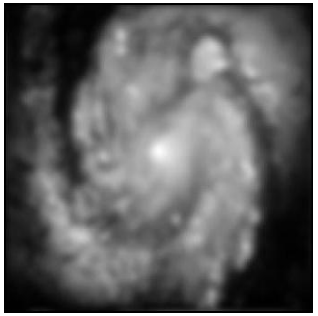
$$R = \langle X^T X \rangle A^T \left(A \langle X^T X \rangle A^T + \langle N^T N \rangle\right)^{-1} = C_X A^T \left(A C_X A^T + C_N\right)^{-1} = \left(A^T C_N^{-1} A + C_X^{-1}\right)^{-1} A^T C_N^{-1}$$

Exercices



- **Filtre optimal**
  - ***Question 4 :*** Dans l'espace de Fourier, les matrices de variance-covariance sont diagonales également et se réduisent aux spectres de puissances. Montrer que le filtre de Wiener ***R*** inverse les basses fréquences et coupe les plus grandes où le bruit domine.

Exercices



Filtre optimal

Solution

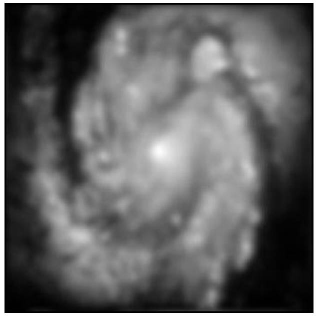
$$R_k = \left(\frac{\lambda_k^2}{C} + \frac{1}{C_0 k^{-\alpha}} \right)^{-1} \frac{\lambda_k}{C} = \frac{\lambda_k}{\lambda_k^2 + \dfrac{C}{C\ k^{-\alpha}}}$$

Pour *k* petit donc $C_0 k^{-\alpha} \gg C$ et $R_k \to \lambda_k^{-1}$

Et pour *k* grand donc $C_0 k^{-\alpha} \ll C$ et $R_k \to 0$

Les très hautes fréquences dominées par le bruit sont coupées, l'image est donc filtrée.

Exercices



- **Filtre optimal**