

1 Filtrage optimal

1.1 Principe d'orthogonalité

Supposons que l'on dispose d'observations $y(t)$ et que l'on veuille élaborer à partir d'elles une estimée *linéaire* d'une quantité d'intérêt $d(t)$ (il peut s'agir d'un *filtrage simple* au cas où l'on aurait $y(t) = x(t) + b(t)$, avec $b(t)$ un bruit additif de caractéristiques statistiques connues, et l'on voudrait estimer $x(t)$, ou encore d'une *prédiction* si l'on veut inférer la valeur $y(t + \tau)$ à partir de la connaissance de l'observation jusqu'au seul instant courant t , etc.).

Dans le cas général, $d(t)$ n'appartient pas nécessairement à l'espace des solutions engendré linéairement par les observations. Il en résulte que, faute de pouvoir répondre exactement à la question posée, il faut se contenter de trouver une solution qui soit la plus proche en un sens donné de ce qui est cherché. Un critère naturel en ce sens est de minimiser l'*erreur quadratique moyenne* entre $d(t)$ et son estimée $\hat{d}(t)$:

$$\mathbb{E}\{[d(t) - \hat{d}(t)]^2\} \rightarrow \min.$$

En s'appuyant sur le fait que l'opérateur d'espérance mathématique définit un produit scalaire permettant d'accéder à une mesure de distance entre les quantités aléatoires considérées, une interprétation géométrique simple du problème de minimisation posé conduit à retenir pour solution l'estimée associée à la projection orthogonale de $d(t)$ sur l'espace des observations. C'est le *principe d'orthogonalité*, stipulant que la solution de filtrage linéaire optimal (à erreur quadratique moyenne minimale) est obtenue en rendant l'erreur orthogonale aux observations au sens du produit scalaire retenu, c'est-à-dire en rendant erreur et observations décorréliées :

$$\mathbb{E}\{[d(t) - \hat{d}(t)]y(v)\} = 0.$$

1.2 Filtrage de Wiener

Soit $d(t)$ la quantité "désirée" et $\hat{d}(t)$ son estimée, élaborée par filtrage linéaire à partir de l'observation $y(t)$:

$$\hat{d}(t) = \int h(t - u) y(u) du. \tag{1}$$

On appelle *filtre de Wiener* le filtre linéaire de réponse $h_*(t)$, optimal au sens où l'erreur associée à l'estimation, c'est-à-dire la différence

$$e(t) = d(t) - \hat{d}(t),$$

est telle que sa puissance est minimisée :

$$P = \mathbb{E}\{e^2(t)\} \rightarrow \min.$$

Par application du principe d'orthogonalité discuté en Sect. 1.1, la solution est obtenue lorsque cette erreur est orthogonale aux observations :

$$\mathbb{E}\{e(t)y(v)\} = 0,$$

d'où l'on déduit, en combinant cette équation et (1) que l'on doit avoir¹ :

$$\gamma_{d,y}(\tau) = \int h(\tau - \theta) \gamma_y(\theta) d\theta.$$

On reconnaît une équation de convolution dans le domaine temporel, dont on sait qu'elle se transforme en produit dans l'espace de Fourier associé des fréquences, conduisant à l'expression du gain complexe du filtre optimal de Wiener donnée par :

$$H_*(f) = \frac{\Gamma_{d,y}(f)}{\Gamma_y(f)}. \quad (2)$$

Du fait que, par construction, l'estimée est orthogonale aux observations, l'erreur quadratique minimale P_* associée au filtre de Wiener se réduit à :

$$P_* = \mathbb{E}\{[d(t) - \hat{d}(t)]d(t)\},$$

soit encore :

$$P_* = \gamma_d(0) - \int h(\tau) \gamma_{d,y}(\tau) d\tau. \quad (3)$$

Passant dans le domaine des fréquences et utilisant la propriété d'isométrie de la transformation de Fourier, on peut ré-écrire cette dernière équation selon

$$P_* = \int \Gamma_d(f) df - \int H^*(f) \Gamma_{d,y}(f) df,$$

¹On convient de noter $\gamma_{x,y}(\tau) := \mathbb{E}\{x(t)y(t - \tau)\}$ et d'adopter la notation simplifiée $\gamma_x(\tau) := \gamma_{x,x}(\tau)$.

soit encore

$$P_* = \int \Gamma_d(f) [1 - |C_{d,y}(f)|^2] df$$

en posant

$$C_{d,y}(f) := \frac{\Gamma_{d,y}(f)}{\sqrt{\Gamma_d(f) \Gamma_y(f)}}.$$

Cette dernière quantité, qui est appelée *cohérence*, vérifie l'inégalité

$$|C_{d,y}(f)| \leq 1,$$

la borne supérieure étant atteinte lorsque $d(t)$ est une filtrée linéaire de $y(t)$. Dans le contexte de filtrage optimal qui nous intéresse ici, on voit donc qu'une cohérence unité a pour conséquence une puissance d'erreur identiquement nulle, ce qui est en accord avec le fait que la quantité désirée appartient à l'espace engendré linéairement à partir des observations.

1.3 Filtrage inverse

Un cas particulièrement important où le filtrage de Wiener peut trouver une application est celui du *filtrage inverse* dans lequel l'observation $y(t)$ est modélisée comme la version déformée d'un signal d'intérêt $x(t)$ par une *fonction d'appareil* $h(t)$ (supposée connue) à laquelle s'ajoute un bruit $b(t)$ que l'on supposera indépendant du signal, centré, stationnaire et de densité spectrale de puissance $\Gamma_b(f)$ connue :

$$y(t) = \int g(t-u) x(u) du + b(t). \quad (4)$$

Ce modèle étant posé, le filtrage inverse consiste à remonter au signal, c'est-à-dire à considérer

$$d(t) = x(t),$$

sur la base des observations $y(t)$ disponibles et des connaissances *a priori* quant à la fonction d'appareil $g(t)$ et au bruit $b(t)$.

La solution générale au problème étant donnée par (2), on évalue dans un premier temps la corrélation croisée entre le signal à estimer et l'observation, ce qui donne :

$$\gamma_{d,y}(\tau) = \int g(\theta - \tau) \gamma_x(\theta) d\theta,$$

équation de corrélation dont la transformée de Fourier fournit directement :

$$\Gamma_{d,y}(f) = G^*(f)\Gamma_x(f).$$

Dans un deuxième temps, on calcule la densité spectrale de puissance de l'observation, ce qui se fait de manière immédiate en utilisant la relation d'entrée-sortie des filtres linéaires et l'indépendance entre signal et bruit :

$$\Gamma_y(f) = |G(f)|^2\Gamma_x(f) + \Gamma_b(f).$$

Reportant ces deux quantités dans (2), on aboutit au résultat cherché qui s'écrit :

$$H_*(t) = \frac{G^*(f)\Gamma_x(f)}{|G(f)|^2\Gamma_x(f) + \Gamma_b(f)} \quad (5)$$

et peut encore se mettre sous la forme :

$$H_*(f) = \frac{1}{G(f)} \frac{1}{1 + \rho^{-1}(f)}$$

en introduisant la quantité :

$$\rho(f) := \frac{|G(f)|^2\Gamma_x(f)}{\Gamma_b(f)}$$

dont l'interprétation physique est celle d'un *rapport signal-sur-bruit*, local en fréquence.

Reportant de la forme explicite de la corrélation croisée entre $d(t) = x(t)$ et $y(t)$, on est conduit à :

$$\begin{aligned} \int h(\tau) \gamma_{d,y}(\tau) d\tau &= \iint h(\tau) g(\theta - \tau) \gamma_x(\theta) d\theta d\tau \\ &= \int H(f) G(f) \Gamma_x(f) df. \end{aligned}$$

En reportant cette équation dans (3) et en remarquant que :

$$\gamma_x(0) = \int \Gamma_x(f) df,$$

on obtient avec (5) :

$$P_* = \int \left(1 - \frac{|G(f)|^2\Gamma_x(f)}{|G(f)|^2\Gamma_x(f) + \Gamma_b(f)} \right) \Gamma_x(f) df,$$

expression que l'on peut simplifier en la ré-écrivaint :

$$P_* = \int \frac{\Gamma_x(f)}{1 + \rho(f)} df.$$

1.3.1 Cas particuliers

On peut déduire des expressions générales précédentes plusieurs cas particuliers, qui en révèlent par exemple les situations limites.

Ainsi, dans le cas où le bruit devient négligeable, c'est-à-dire lorsque le rapport signal-sur-bruit $\rho(f)$ tend vers l'infini, on obtient comme attendu :

$$\lim_{\rho(f) \rightarrow \infty} H_*(f) = \frac{1}{G(f)}$$

et

$$\lim_{\rho(f) \rightarrow \infty} P_* = 0.$$

L'interprétation en est qu'en l'absence de bruit, la réponse fréquentielle du filtre inverse n'est autre que l'inverse de la fonction d'appareil, la connaissance supposée parfaite de cette dernière conduisant en outre à une erreur nulle.

À l'inverse, si l'on suppose maintenant que le bruit devient prépondérant au sens où le rapport signal-sur-bruit tend vers zéro, on obtient immédiatement que :

$$\lim_{\rho(f) \rightarrow 0} H_*(f) = 0$$

et

$$\lim_{\rho(f) \rightarrow 0} P_* = \gamma_x(0).$$

Dans ce cas où l'information utile portée par le signal est totalement noyée dans le bruit d'observation, l'effet additionnel de sa modification par l'appareil de mesure devient négligeable et la meilleure estimation se réduit à la valeur moyenne du bruit (nulle par hypothèse), avec une erreur quadratique s'identifiant à la variance de ce bruit (valeur de la corrélation à l'origine).

On peut enfin noter que si l'on s'affranchit de l'effet d'une éventuelle fonction d'appareil et que l'on s'intéresse au seul *débruitage* d'un signal, on se place dans le cas simplifié :

$$G(f) = 1 \Rightarrow H_*(f) = \frac{\Gamma_x(f)}{\Gamma_x(f) + \Gamma_b(f)} \quad (6)$$

pour lequel les remarques précédentes relatives aux cas limites de rapports signal-sur-bruit nul ou infini continuent bien sûr à s'appliquer.

1.3.2 Exemples

Exemple 1. `filter_chirp_WGN.m` (Filtrage simple) — On se place dans le cas $G(f) = 1$, mais avec un signal inconnu dans un bruit inconnu. Dans sa plus grande généralité, c'est évidemment un problème mal posé mais, dans le cas où l'on dispose d'une "observation bruit seul", on peut ré-écrire (6) selon

$$H_*(f) = 1 - \frac{\Gamma_b(f)}{\Gamma_y(f)}$$

et construire une estimée de cette réponse en estimant d'une part la densité spectrale de puissance globale de l'observation et d'autre part celle du bruit à partir d'un intervalle temporel supposé "bruit seul".

Exemple 2. `ex_deconv.m` (Déconvolution) — On se place ici dans le cas d'images pour lesquelles la fonction d'appareil (flou gaussien, flou rectangulaire, flou de bougé en translation unidirectionnelle) est supposée parfaitement connue, mais susceptible de posséder des zéros (conduisant à une divergence du filtre inverse) ou de très faibles valeurs (conduisant à une grande instabilité de la solution en présence de bruit). L'objet de cet exemple n'est pas de fournir une solution solide à cette question mais de sensibiliser au problème et de voir l'effet *régularisant* de l'introduction d'un terme correctif ε dans le filtre inverse :

$$H_*(f) = \frac{1}{G(f)} \rightarrow \frac{1}{G(f) + \varepsilon/G^*(f)},$$

s'apparentant à la prise en compte d'un "bruit équivalent" de rapport signal-sur-bruit

$$\rho(f) = \frac{|G(f)|^2}{\varepsilon}.$$