FILTRAGE (C2, J4, N3)
(25 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

La notion de **filtrage** est une transformation généralement effectuée sur un processus stochastique (ou sur une série temporelle) (cf aussi **transformation des données**, **changement de variable aléatoire**).

Un filtrage est utilisé pour une **analyse dans le domaine des états** (cf **espace des états**) ou pour une **analyse dans le domaine du temps**, ou **analyse dans le domaine des fréquences** (cf **espace du temps**).

On suppose définies les notions de processus et d'**intégrale** à valeurs dans des espaces de BANACH.

(i) Soit $\{(\Omega, \mathscr{T}, P), (\mathscr{X}, \mathscr{B}), (X_t)_{t \in T}\}$ un **processus stochastique** tq $\mathscr{X}$ est un **espace de BANACH**, muni de sa **tribu borélienne** $\mathscr{B}$, et T un **groupe** additif abélien (eg T = **Z** ou T = **R**).

On appelle **filtrage**, ou **filtre**, (**non linéaire**) de la **famille** $X = (X_t)_{t \in T}$, ou du processus X, toute opération $\Phi : \mathscr{X}^T \mapsto \mathscr{X}^T$ définie par :

$$(1) \qquad X \in \mathscr{X}^T \mapsto Y = \Phi(X), \quad \text{avec} \quad Y_t = \int_T \phi(X_{t+\tau})\, d\mu(\tau),$$

pour tout $t \in T$ tq $t + \tau \in T$, dans laquelle $\phi : \mathscr{X} \mapsto \mathscr{X}$ est une **application mesurable** donnée et μ une **mesure abstraite** positive finie sur $\mathscr{B}_T$ (**tribu de parties** de T).

On appelle aussi **filtrage**, ou parfois **noyau du filtre**, l'application ϕ qui sert à définir le filtrage Φ.

Lorsque ϕ est une **application linéaire** (cf **homomorphisme**), Φ est aussi linéaire, et l'on définit ainsi un ensemble très important de filtres, les **filtres linéaires**. On dit alors que ϕ (ou Φ) est un **filtrage linéaire** de X.

(ii) Deux exemples classiques de filtres (non linéaires) sont les suivants :

(a) si T = **Z** et si μ est une **mesure discrète** (ie de la forme $\mu = \Sigma_{n \in \mathbf{Z}}\, \alpha_n\, \delta_n$) sur $\mathscr{B}_{\mathbf{Z}}$, on obtient :

$$(2) \qquad Y_t = \Sigma_{\tau \in \mathbf{Z}}\ \alpha_\tau \,.\, \phi(X_{t+\tau}), \qquad \forall\, t \in \mathbf{Z}.$$

En particulier, si $\phi = id_{\mathscr{X}}$, $m > 0$, $n > 0$ et $\alpha_\tau = 0$ pour tout $\tau \notin Z_{mn} = \{- m,..., - 1, 0, + 1,..., n\}$, on définit un **filtre de moyenne mobile** de X (cf **moyenne mobile**, **processus de moyenne mobile**) ;

(b) si T = **R** et si μ est la **mesure de LEBESGUE** sur $\mathscr{B}(\mathbf{R})$, on obtient :

(3) $Y_t = \int_T \phi (X_{t+\tau}) \, d\tau, \quad \forall t \in \mathbf{R}.$

On adopte parfois une définition du filtre différente de (1), ie :

(4) $Y_t = \int_T \phi (X_{t-\tau}) \, d\mu (\tau), \quad \forall t \in T \text{ tq } t - \tau \in T.$

(iii) En général, μ est une **mesure de probabilité** sur $\mathscr{B}_T$ définie à l'aide d'une **fonction de poids**, ou **fonction de pondération** (ie la **densité de probabilité** associée).

(iv) Lorsque le filtre est une moyenne mobile symétrique (ie n = m, avec les notations précédentes) et que X est un **processus stationnaire en covariance** et de spectre (cf **densité spectrale**) associé f, le processus image Y admet pour spectre la densité spectrale g tq :

(5) $g (\omega) = a (\omega) . f (\omega), \quad \forall \omega \in [- \pi, + \pi],$

avec $a (\omega) = \alpha (\omega) \, \overline{\alpha} (\omega)$, où $\alpha (\omega) = \Sigma_{\tau=-m}^{m} \; \alpha_\tau \, . \, e^{i \omega \tau}$ et $\overline{z}$ désigne le nombre conjugué de $z \in \mathbf{C}$.

(v) Si T = **R** et si $X = (X_t)_{t \in T}$ est un processus scalaire complexe (ie si $\mathscr{X} = \mathbf{C}$), alors le **processus filtré** (processus de moyenne mobile d'ordre $p > 1$) :

(6) $Y_t = \Sigma_{j=0}^{p} \alpha_j . X_{t-j} = \alpha (L) X_t ,$

dans lequel L est l'**opérateur retard** $X_t \mapsto L \, X_t = X_{t-1}$ et $\alpha (L) = \Sigma_{j=0}^{p} \alpha_j L^j$ est un polynôme formel, admet pour densité spectrale :

(7) $g (\omega) = |\alpha (e^{i\omega})|^2 . f (\omega), \quad \forall \omega \in [- \pi, + \pi].$

(vi) Le **problème de filtrage** consiste principalement à déterminer un filtre optimal Φ^* permettant de séparer un **signal** pur inobservable X* d'un autre processus U (généralement appelé **bruit**) perturbateur de X (cf **théorie du signal**). Le processus observé dépend donc de X* et de U, eg X = f (X*, U). L'**estimation** de Φ est donc dépendante de celle de X* : l'application de Φ^* à X doit ainsi conduire à X*.

Un exemple important est celui du **filtre de KALMAN**.

(vii) Les notions précédentes possèdent des analogues directes dans le cas d'une **série temporelle**. Ainsi, on considère une série chronologique en **temps** discret $x = (x_t)_{t \in \mathbf{N}}$ à valeurs dans un **espace vectoriel** $\mathscr{X}$, une **partie** donnée $J \in \mathscr{P}(\mathbf{N})$ (avec, le plus souvent, Card $J < +\infty$) et une **suite** double $(A (t, j))_{(t, j)}$ constituée d'**opérateurs linéaires** dans $\mathscr{X}$ (ie $A (t, j) \in \mathscr{L}(\mathscr{X})$), avec $(t, j) \in \mathbf{N} \times J$.

On appelle **filtre linéaire** une opération Φ, définie dans l'espace $\mathscr{X}^{\mathbf{N}}$ des séries temporelles à valeurs dans $\mathscr{X}$ selon :

(8) $\Phi : x \mapsto y$, avec $y_t = \Sigma_{j \in J} A(t, j) \, x_{t+j}$, $\forall \, t \in \mathbf{N}$.

L'opération Φ est évidemment linéaire dans $\mathscr{X}^{\mathbf{N}}$.

Un exemple de filtre linéaire couramment utilisé est le **filtre de moyenne mobile** (défini comme en (2)).

(viii) En pratique, un **filtrage** sert souvent aussi :

(a) à « lisser » un processus ou une série temporelle, ie à en déduire un processus (ou une série) plus « simple » ou plus « régulièr(e) » ;

(b) à extraire la **tendance** ou les composantes harmoniques d'un processus (ou d'une série temporelle) (cf **analyse harmonique**, **analyse spectrale**).