Estimation du canal avec séquence d'apprentissage

Philippe Ciblat

École Nationale Supérieure des Télécommunications, Paris, France

Modèle en bande de base

Signal reçu en bande de base

$$y_a(t) = \sum_{k\in\mathbb{Z}} s_k h_a(t - kT_s) + b_a(t)$$

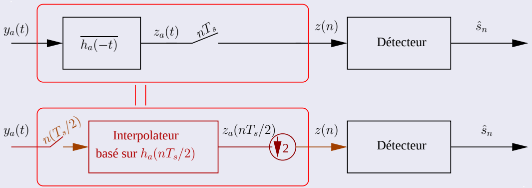
avec

- $\{s_k\}_{k\in\mathbb{Z}}$: symboles d'une constellation quelconque
- $b_a(t)$: bruit additif gaussien $\sim \mathcal{CN}(0, 2N_0)$
- $h_a(t)$: un filtre résultant d'un filtre de mise en forme $g_a(t)$, d'un canal de propagation $c_a(t)$ de type multitrajet. Ce filtre a pour support temporel $[0, LT_s]$

$\rightsquigarrow$ *Le canal $h_a(t)$ est, en pratique, inconnu du récepteur*

Récepteur avec canal $h_a(t)$ connu

Schéma de principe



Remarque :

- Le filtre de mise en forme admet une bande de la forme $[-(1+\rho)/(2T_s), (1+\rho)/(2T_s)]$ avec ρ le facteur d'excès de bande
- Période d'échantillonnage $T_e = T_s/2$ vérifie le théorème de Shannon

Récepteur avec canal $h_a(t)$ inconnu

Schéma de principe (avec filtre adapté)

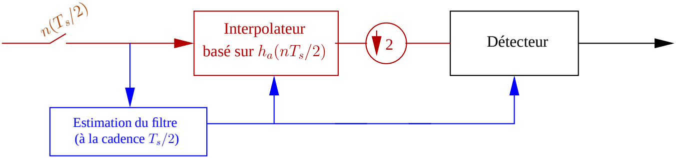
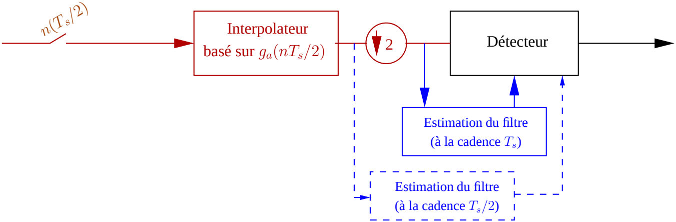


Schéma de principe (sans filtre adapté)



Formulation mathématique

Signal à temps discret (suréchantillonné)

$$\tilde{y}(n) = y_a(nT_e) = y_a(nT_s/2) = \sum_{k\in\mathbb{Z}} s_k h_a(nT_s/2 - kT_s) + \tilde{b}(n)$$

Modèle équivalent

$$\left\{\begin{array}{lclcl} y_1(n) & = & y_a(nT_s) & = & h_1 \star s_n + b_1(n) \\ y_2(n) & = & y_a(nT_s + T_s/2) & = & h_2 \star s_n + b_2(n) \end{array}\right.$$

avec

- $h_1(n) = h_a(nT_s)$
- $h_2(n) = h_a(nT_s + T_s/2)$

Modèle formel

$$y(n) = h \star s_n + b(n)$$

avec h un filtre de degré L (sur chaque voie)

Méthode d'estimation

Trois modes :

- Avec séquence d'apprentissage
  $\rightsquigarrow$ Data-aided (DA) ou supervisé
- Sans séquence d'apprentissage
  $\rightsquigarrow$ Non-Data-aided (NDA) ou autodidacte/aveugle
- Avec retour de décision
  $\rightsquigarrow$ Decision-Directed (DD)

Objectif

Estimer h à la donnée de $y(n)$ et s_n

Formulation matricielle

$$\mathbf{y}_N = \mathbf{S}_N \mathbf{h} + \mathbf{b}_N$$

avec

- $\mathbf{y}_N = [y(0), \cdots, y(N-1)]^{\mathrm{T}}$
- N nombre d'observations (temps d'observation $= [0, NT_s[$)
- $[\mathbf{0}_{1,L}, s_0, \cdots, s_{N-1}]^{\mathrm{T}}$: séquence d'apprentissage
- $\mathbf{S}_N$ une matrice $N \times (L+1)$ définie comme suit

$$\mathbf{S}_N = \begin{bmatrix} s_0 & s_{-1} & \dots & s_{-L} \\ s_1 & \ddots & \ddots & \\ \vdots & & & \\ s_{N-1} & s_{N-2} & \dots & s_{N-1-L} \end{bmatrix}$$

- $\mathbf{h} = [h(0), \cdots, h(L)]^{\mathrm{T}}$
- $\mathbf{b}_N = [b(0), \cdots, b(N-1)]^{\mathrm{T}}$: bruit blanc gaussien

Estimateur du maximum de vraisemblance

On a

$$p(\mathbf{y}_N|\mathbf{h}) \propto \exp\left\{-\frac{\|\mathbf{y}_N - \mathbf{S}_N\mathbf{h}\|^2}{2N_0}\right\}$$

d'où

$$\hat{\mathbf{h}}_N = \arg\min_{\mathbf{h}} \|\mathbf{y}_N - \mathbf{S}_N\mathbf{h}\|^2$$

Résultat

$$\hat{\mathbf{h}}_N = \mathbf{S}_N^{\#}\mathbf{y}_N = \left(\mathbf{S}_N^{\mathrm{H}}\mathbf{S}_N\right)^{-1}\mathbf{S}_N^{\mathrm{H}}\mathbf{y}_N$$

avec $\mathbf{S}_N^{\#}$ la pseudo-inverse de $\mathbf{S}_N$

Performances asymptotiques (I)

Un résultat préliminaire

- $\{s_n\}_n$ réalisation d'une séquence pseudo-aléatoire stationnaire
- $r_s(\tau) = \mathbb{E}[s_{n+\tau}\overline{s_n}]$ la fonction d'autocorrélation
- $\mathbf{T}_N = \frac{1}{N}\mathbf{S}_N^{\mathrm{H}}\mathbf{S}_N$

Résultat fondamental

$$t_N(k,l) = \frac{1}{N}\sum_{n=0}^{N-1} \overline{s_{n-k}}s_{n-l} \overset{p.s.}{\to} \mathbb{E}[\overline{s_{n-k}}s_{n-l}] = r_s(k-l)$$

Considérons $\mathbf{R}_s = (r_s(k-l))_{0\leq k,l\leq L}$ la matrice (Toeplitz) de corrélation de taille $(L+1)\times(L+1)$ du processus $\{s_n\}_n$

$$\frac{1}{N}\mathbf{S}_N^{\mathrm{H}}\mathbf{S}_N \overset{p.s.}{\to} \mathbf{R}_s$$

Performances asymptotiques (II)

Résultat

- Estimateur sans biais : $\mathbb{E}[\hat{\mathbf{h}}_N] = \mathbf{h}$
- Estimateur consistant

$$\hat{\mathbf{h}}_N \stackrel{p.s.}{\rightarrow} \mathbf{h}$$

- Estimateur asymptotique normal (avec une vitesse en $1/N$)

$$N^{1/2}(\hat{\mathbf{h}}_N - \mathbf{h}) \stackrel{\mathcal{L}}{\rightarrow} \mathcal{CN}\left(0, \mathbf{\Gamma}\right)$$

avec

$$\mathbf{\Gamma} = 2N_0 \left(\mathbf{R}_s\right)^{-1}$$

Remarque :

$$\mathrm{MSE} = \mathbb{E}\left[\|\hat{\mathbf{h}}_N - \mathbf{h}\|^2\right] \approx \frac{\mathrm{trace}(\mathbf{\Gamma})}{N}$$

Preuve

$$
\begin{aligned}
\hat{\mathbf{h}}_N &= \left(\mathbf{S}_N^{\mathrm{H}}\mathbf{S}_N\right)^{-1}\mathbf{S}_N^{\mathrm{H}}\mathbf{y}_N \\
&= \mathbf{h} + \left(\mathbf{S}_N^{\mathrm{H}}\mathbf{S}_N\right)^{-1}\mathbf{S}_N^{\mathrm{H}}\mathbf{b}_N \\
&\Rightarrow \text{sans biais} \\
&= \mathbf{h} + \left(\frac{1}{N}\mathbf{S}_N^{\mathrm{H}}\mathbf{S}_N\right)^{-1}\frac{1}{N}\mathbf{S}_N^{\mathrm{H}}\mathbf{b}_N \\
&\Rightarrow \text{consistant} \\
N^{1/2}(\hat{\mathbf{h}}_N - \mathbf{h}) &= \left(\frac{1}{N}\mathbf{S}_N^{\mathrm{H}}\mathbf{S}_N\right)^{-1}\left(\frac{1}{\sqrt{N}}\mathbf{S}_N^{\mathrm{H}}\mathbf{b}_N\right) \\
&\Rightarrow \text{asymptotiquement normal} \\
N\mathbb{E}[(\hat{\mathbf{h}}_N - \mathbf{h})(\hat{\mathbf{h}}_N - \mathbf{h})^{\mathrm{H}}] &= 2N_0\left(\frac{1}{N}\mathbf{S}_N^{\mathrm{H}}\mathbf{S}_N\right)^{-1} \\
&\Rightarrow \text{de covariance asymptotique } \mathbf{\Gamma}
\end{aligned}
$$

Séquence d'apprentissage optimale

- Les performances dépendent donc de la couleur de la séquence d'apprentissage
- Question : quelle est la couleur optimale ?

Problème d'optimisation :

$$\mathbf{R}_{\text{opt.}} = \arg \min_{\substack{\mathbf{R}_S \\ \text{trace}(\mathbf{R}_S)=(L+1)P_0}} \text{trace}(\mathbf{R}_s^{-1})$$

Résultat

La matrice $\mathbf{R}_s = P_0 \mathbf{Id}_{L+1}$ rend la MSE minimale

Remarque :

- La séquence d'apprentissage optimale est *blanche*
- L'estimateur ML devient le simple corrélateur $\hat{\mathbf{h}}_N = \frac{1}{NP_0} \mathbf{S}_N^{\mathrm{H}} \mathbf{y}_N$

Preuve

- La matrice $\mathbf{R}_s$ est hermitienne et donc diagonalisable
- Soit $\mathbf{R}_s = \mathbf{P}^{-1}\boldsymbol{\Lambda}\mathbf{P}$ avec $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_0, \ldots, \lambda_L)$
- On a

$$\mathbf{R}_{\mathrm{opt.}} = \arg \min_{\substack{\lambda_0, \ldots, \lambda_L \\ \sum_{\ell=0}^{L} \lambda_\ell = (L+1)P_0}} \sum_{\ell=0}^{L} \frac{1}{\lambda_\ell}$$

- La fonction $f(x) = 1/x$ étant convexe, on a

$$f\left(\frac{1}{L+1}\sum_{\ell=0}^{L}\lambda_\ell\right) \leq \frac{1}{L+1}\sum_{\ell=0}^{L} f(\lambda_\ell) \Rightarrow \sum_{\ell=0}^{L}\frac{1}{\lambda_\ell} \geq \frac{L+1}{P_0}$$

- La borne est atteinte si

$$\lambda_\ell = P_0$$

Borne de Cramer-Rao

Résultat

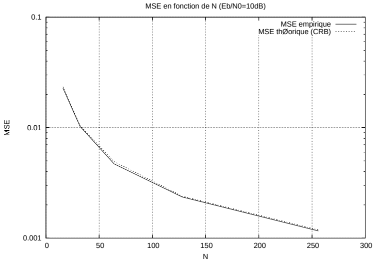
$$\mathrm{CRB}_N(\mathbf{h}) = \frac{2N_0}{N}\left(\frac{1}{N}\mathbf{S}_N^{\mathrm{H}}\mathbf{S}_N\right)^{-1}$$

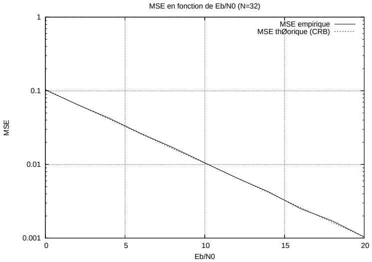
Remarque :
L'estimateur ML est donc efficace

Exemple numérique : EQM (*MSE*)

On a

- $\mathbf{h} = [-0.40825, 0.81650, 0.40825]^{\mathrm{T}}$
- Séquence blanche



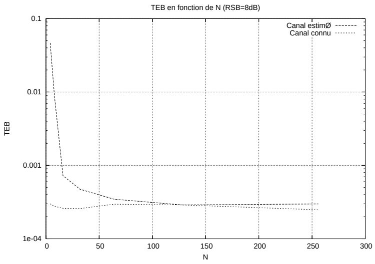


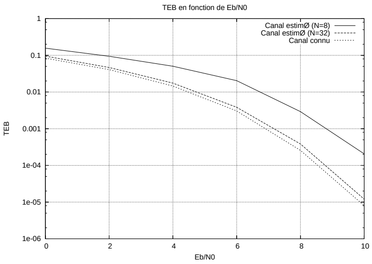
Exemple : $N = 32$ pour une trame de longueur 128 dans le GSM

Exemple numérique : TEB (*BER*)

On a

- $\mathbf{h} = [-0.40825, 0.81650, 0.40825]^{\mathrm{T}}$
- Séquence blanche
- Egaliseur de Wiener
- Constellation d'amplitude à deux états (MDA2/BPSK)





Bibliographie

- S. Crozier, D. Falconer et S. Mahmoud, "Least sum of squared errors channel estimation", IEE Proceedings on Radar and Signal Processing, vol. 138, no. 4, pp. 371-378, Août 1991.
- C. Tellambura, "Optimal sequences for channel estimation using FFT techniques", IEEE Trans. on Communications, vol. 47, no. 2, pp 230-238, Février 1999.
- M. Morelli et U. Mengali, "Carrier-frequency estimation for transmissions over selective channels", IEEE Trans. on Communications, vol. 48, pp. 1580-1589, Septembre 2000.
- P. Ciblat et L. Vandendorpe, "On the Maximum-Likelihood based data-aided frequency offset and channel estimates", EURASIP European Signal Processing Conference (EUSIPCO), vol. 1, pp. 627-630, Toulouse (France), Septembre 2002.
- P. Stoica et O. Besson, "Training sequence design for frequency offset and frequency-selective channel estimation", IEEE Trans. on Communications, vol. 51, no. 11, pp 1910-1917, Novembre 2003.