COLLOQUE NATIONAL SUR LE TRAITEMENT DU SIGNAL
ET SES APPLICATIONS

NICE du 16 au 21 JUIN 75

APPLICATION EN DETECTION D'UN CRITERE DE MINIMALISATION DE DISTANCE D'OBSERVATION

Pierre-Yves ARQUES

UNIVERSITE DE RENNES, Laboratoire de Traitement du Signal, Avenue du Général Leclerc, 35031-RENNES CEDEX

RESUME

On considère un modèle d'étude de problèmes de décision utilisant un critère de minimalisation de distance dans l'espace d'observation. Il recouvre en particulier le modèle d'estimation selon les moindres carrés pondérés. On considère l'application aux problèmes de classification et de détection ; on en déduit des structures de systèmes, optimaux au sens précédent, pour la détection de certains types de signaux perturbés par des bruits : signal certain complètement connu, ou de phase inconnue, ou d'amplitude inconnue, ou de direction inconnue.

SUMMARY

A model for decision problems, using a minimum distance criterion, is presented ; the method of weighted least square in estimation is a particular case. The application to m-ary detection is studied ; optimum structures, in this sense, are obtained for the detection of signal with completely known form, or with unknown phase, or with unknown amplitude, or with unknown direction.

APPLICATION EN DETECTION D'UN CRITERE DE MINIMALISATION DE DISTANCE D'OBSERVATION

1 - GENERALITES

Un problème de décision est ici un problème de choix optimal concernant une propriété d'un paramètre θ, à partir d'une réalisation x d'un élément aléatoire X dont la loi de probabilité dépend du paramètre θ. Un modèle permettant de schématiser un problème de décision peut être constitué (figure 1) à partir de :

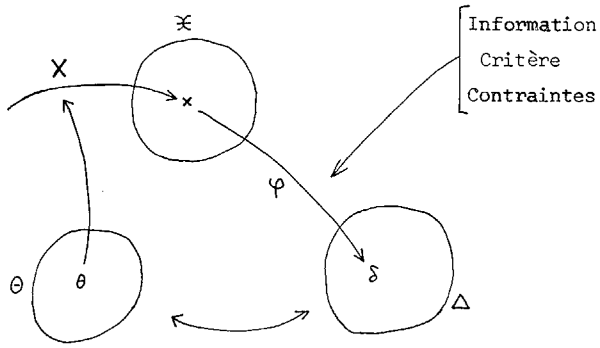

FIGURE 1

- une structure décisionnelle : c'est la donnée d'un espace Θ des paramètres θ, d'un espace $\mathfrak{X}$ des observations x c'est-à-dire des réalisations d'un élément aléatoire X, d'un espace Δ des décisions δ ; Δ est en liaison avec Θ ou $\mathcal{P}(\Theta)$, ensemble des parties de Θ ;
- un ensemble Φ de fonctions de décision pures, applications de $\mathfrak{X}$ dans Δ ;
- une information concernant les propriétés de l'observation aléatoire X et du paramètre θ ;
- un critère d'optimalisation ;

Une optimalisation sera dite à structure imposée, ou à structure libre selon que l'ensemble Φ est limité ou non par une "contrainte" [1 à 3] .

Les problèmes d'estimation, de détection, de classification apparaissent comme des cas particuliers. Leur étude montre qu'il est possible de développer différents types de modèles. Le choix du modèle peut se faire en fonction de l'information disponible. Les détecteurs à structure libre, par exemple, sont généralement construits à partir d'un modèle bayesien ou d'un modèle à critère d'erreur de décision (Neyman-Pearson). Ces modèles nécessitent pour être développés une information importante : l'observateur doit connaître la loi de probabilité de l'observation sous chacune des hypothèses considérées. Il est possible de construire d'autres modèles, généralement à structure imposée, à partir d'un critère de contraste dans un espace de transition entre les espaces $\mathfrak{X}$ et Δ ; en détection c'est un critère de rapport signal sur bruit, nécessitant moins d'information que les modèles précédents (par exemple une connaissance au second ordre).

On peut ainsi proposer un modèle particulier de décision à critère de contraste, nécessitant une information réduite, et s'appuyant sur l'utilisation de distances dans l'espace d'observation. Un cas particulier d'un tel modèle est connu en estimation sous la forme "moindres carrés pondérés". Ces modèles à critère de distance d'observation, s'ils semblent pouvoir être aisément mis en oeuvre dans certaines catégories de problèmes de décision, permettent plus difficilement l'évaluation des performances du système obtenu. Ils permettent de retrouver différentes structures classiques en détection. Ils semblent surtout présenter l'intérêt de s'adapter à une information réduite ; le manque d'information est compensé par une spécification arbitraire de la structure et une moindre élaboration du critère de décision.

2 - MODELES A CRITERE DE CONTRASTE

Un modèle à critère de contraste, résulte de la donnée ou du choix (figure 2) :

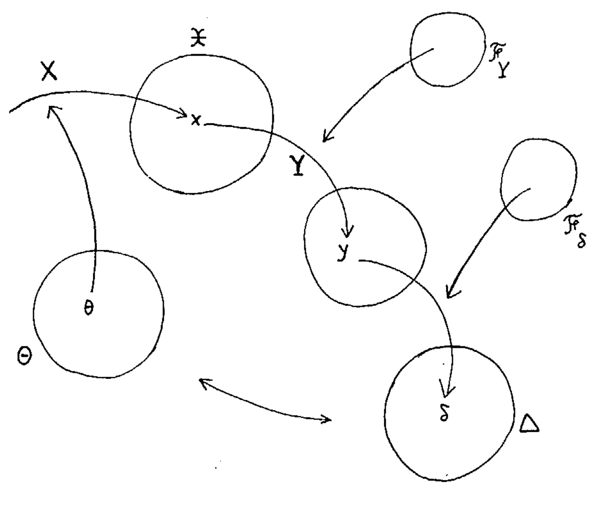

FIGURE 2

- d'une structure décisionnelle $(\Theta,\mathfrak{X},\Delta)$;
- d'un espace de transition $\mathcal{Y}$, d'une famille $\mathcal{F}_Y$ d'applications Y de $\mathfrak{X}$ dans $\mathcal{Y}$, et d'une famille $\mathcal{F}_\delta$ d'applications de $\mathcal{Y}$ dans Δ ;

- d'un critère de contraste sur $\mathcal{Y}$ et d'un critère de décision sur Δ .

Une fonction de décision possible est représentée par le produit de composition de deux applications successives par transition dans l'espace $\mathcal{Y}$. Le choix et l'évaluation d'une fonction de décision φ , se fait d'abord au niveau de $\mathcal{Y}$ par choix dans $\mathcal{F}_Y$ au moyen du critère de contraste sur $\mathcal{Y}$, ensuite au niveau de Δ par choix dans $\mathcal{F}_\delta$ au moyen du critère de décision sur Δ . Les familles $\mathcal{F}_Y$ et $\mathcal{F}_\delta$ sont en général restreintes selon les contraintes ou l'information a priori disponible (optimalisation à structure imposée).

3 - MODELES A CRITERE DE DISTANCE DANS L'ESPACE D'OBSERVATION

Plus particulièrement un modèle à critère de distance d'observation résulte (les différentes hypothèses doivent être compatibles) de :

- la donnée d'une structure décisionnelle $(\Theta, \mathcal{X}, \Delta)$;
- l'existence d'une loi de probabilité $P_{X|\theta}$ de l'observation X conditionnelle au paramètre θ et la donnée d'une certaine information sur cette loi ;
- le choix ou la donnée d'une famille de distances d sur $\mathcal{X}$ et d'une famille de références $\overline{x}$ de l'observation x pour chaque sous-ensemble, ou "hypothèse" H, de Θ correspondant à une décision δ possible de Δ ; ainsi que d'une famille $\mathcal{F}_\delta$ d'applications, dans l'espace Δ, de l'espace $\mathcal{Y}$ des valeurs numériques associées à chaque observation par la distance pour chaque hypothèse possible ; on a $\mathcal{Y} = (\mathbb{R}^+)^{\text{card } \Delta}$;
- le choix ou la donnée d'un critère de contraste sur $\mathcal{Y}$ (par exemple un rapport signal sur bruit) et d'un critère de décision sur Δ .

Le choix d'une fonction de décision φ optimale se fait en deux temps :

- Choix, pour chaque hypothèse H possible, au moyen du critère de contraste, dans les familles correspondantes, d'une distance d_H sur $\mathcal{X}$ et d'une référence $\overline{x}(H)$, permettant de faire correspondre à chaque observation x sa distance, $d_H(x,\overline{x}(H))$, en fonction de H à la référence $\overline{x}(H)$.
- Choix, pour chaque x, dans la famille $\mathcal{F}_\delta$, au moyen du critère de décision, d'une décision δ à partir de l'ensemble des valeurs $d_H(x,\overline{x}(H))$.

La famille $\mathcal{F}_\delta$ d'applications de $\mathcal{Y}$ dans Δ sera ici définie par la propriété de faire correspondre à y la décision δ "choix de H", telle que

$$\forall x \in \mathcal{X}, \forall H' , \varkappa_H + d_H^a(x,\overline{x}(H)) \leq \varkappa_{H'} + d_{H'}^a(x,\overline{x}(H'))$$

(a et les $\varkappa_H$ sont des constantes à fixer ou à optimaliser au moyen du critère de décision). On dispose pour adapter le modèle au problème posé du choix conjoint des références observation $\overline{x}(H)$, des distances d_H, et des quantités $\varkappa_H$ et a .

Ce modèle utilisé en estimation généralise l'estimation selon les moindres carrés :

$$\hat{\theta}(x) : \forall x \in \mathcal{X}, \forall \theta \in \Theta , d_{\hat{\theta}}^a(x,\overline{x}(\hat{\theta})) + \varkappa_{\hat{\theta}} \leq d_\theta^a(x,\overline{x}(\theta)) + \varkappa_\theta$$

On peut choisir d_θ indépendante de θ ou telle qu'elle rend maximale pour tout θ un critère de contraste faisant intervenir les distances entre $\overline{x}(\theta)$ et x, selon que x est engendré ou non par θ . Si l'espérance mathématique de X conditionnelle à θ, est, en tant que fonction de θ, connue, on peut choisir pour référence $\overline{x}$ de l'observation en fonction du paramètre :

$$\overline{x} = E_{X|\theta}\{X\}$$

En classification, pour (m+1) hypothèses $H_o, H_1, \ldots, H_m$, la règle de décision est la suivante :

Décision $\delta_j(x)$ "choix de H_j pour x", $\forall j \in \{0,\ldots,m\}$, si $x \in D_j$:

$$D_j = \{x : x \in \mathcal{X}; \forall \ell \in \{0,\ldots,m\}, \ell \neq j ; d_j^a(x,\overline{x}_j) + \varkappa_j \leq d_\ell^a(x,\overline{x}_\ell) + \varkappa_\ell\}$$

On peut par exemple choisir pour une hypothèse simple ou composée H_j :

$$\overline{x}_j = E_{X|H_j}\{X\}$$

ou encore pour une hypothèse composée H_j, si l'on peut déterminer un estimateur $\hat{\theta}(x)$ du paramètre θ, sous H_j, à partir de l'observation x :

$$\overline{x}_j = E_{X|\hat{\theta}(x)}\{X\}$$

4 - APPLICATION A UN SIGNAL PERTURBE PAR UN BRUIT

4-1-

On se restreint dans la suite à préciser quelques uns des résultats obtenus dans le cas particulier suivant :

- l'observation est une tranche de réalisation v(t) d'un signal aléatoire vectoriel V(t), à n_e composantes sur $[t_D-T, t_D]$ (V et v sont des $(n_e,1)$-matrices, $\tilde{V}$ et $\tilde{v}$ sont leurs transposées) ;

- on utilise le carré (a=2) de la distance associée à la norme pondérée :

$$||x|| = \left[\int_{t_D-T}^{t_D} \tilde{v}(t)G(t)v(t)dt\right]^{1/2} \quad ;$$

la (n_e,n_e)-matrice G(t) est symétrique, définie strictement positive.

On peut étudier de manière totalement analogue le cas discret.

4-2-

En estimation d'un paramètre θ, pour une observation constituée d'un signal déterministe vectoriel S dépendant du paramètre inconnu vectoriel θ, perturbé par un bruit parasite vectoriel additif B, on a :

$$V(t) = B(t) + S(t,\theta)$$

et on choisit une référence observation de la forme :

$$\overline{x} = \overline{B}(t) + \overline{S}(t,\theta)$$

L'estimateur est tel que :

$$\hat{\theta}_{MC} \Rightarrow \inf_{\theta}\left\{\int_{t_D-T}^{t_D}\left[\tilde{v}(t)-\tilde{\overline{B}}(t)-\tilde{\overline{S}}(t,\theta)\right]G(t)\left[v(t)-\overline{B}(t)-\overline{S}(t,\theta)\right]dt + \varkappa_\theta\right\}$$

•• On considère une pondération G(t) indépendante de θ : on retrouve pour $\varkappa_\theta = 0$, pour un bruit blanc B(t) de valeur moyenne E{B(t)} connue, pour un signal déterministe S(t) dépendant du seul paramètre inconnu θ, pour une référence bruit $\overline{B}(t) = E\{B(t)\}$ et signal $\overline{S}(t) = S(t,\theta)$, l'estimateur classique selon les moindres carrés.

•• On considère une pondération $G_\theta(t)$ dépendant de θ : on montre dans [3] que, sous certaines hypothèses, l'on peut optimaliser le modèle en déterminant la matrice $G_\theta(t)$ maximalisant un rapport signal sur bruit construit sur d_θ^2.

4-3-

En classification des m+1 hypothèses H_i correspondant à une observation engendrée par un signal vectoriel $S_{(i)}$ parmi m possibles, perturbé par un bruit vectoriel additif B, ou à une absence de signal, on a:

$$H_i \ , \forall i \in \{1,..,m\} : V(t) = B(t) + S_{(i)}(t)$$
$$H_o \qquad : V(t) = B(t)$$

On choisit une référence observation de la forme :

$$\forall\, i\in\{0,..,m\} \ , \ \overline{x}_i = \overline{B}(t) + \overline{S}_{(i)}(t)$$

avec $\qquad \overline{S}_{(0)}(t) = 0$

La règle de décision est alors :

Décision $\delta_j, \forall j\in\{0,..,m\}$, si $\forall \ell\in\{0,..,m\}, \ell\neq j$:

$$\int_{t_D-T}^{t_D}\left[\tilde{v}(t)-\tilde{\overline{B}}(t)-\tilde{\overline{S}}_{(j)}(t)\right]G_j(t)\left[v(t)-\overline{B}(t)-\overline{S}_{(j)}(t)\right]dt+\varkappa_j$$

$$\leqslant \int_{t_D-T}^{t_D}\left[\tilde{v}(t)-\tilde{\overline{B}}(t)-\tilde{\overline{S}}_{(\ell)}(t)\right]G_\ell(t)\left[v(t)-\overline{B}(t)-\overline{S}_{(\ell)}(t)\right]dt+\varkappa_\ell$$

On examine en conséquence quelques résultats particuliers obtenus en détection de présence (m=1).

5 - CAS DE LA DETECTION

5-1-

En détection de présence la règle de décision précédente se ramène à la comparaison, avec $\varkappa = \varkappa_1 - \varkappa_o$, de la quantité :

$$w(t_D) = \int_{t_D-T}^{t_D}\left[(\tilde{v}(t)-2\tilde{\overline{B}}(t))(G_o(t)-G_1(t))+2\tilde{\overline{S}}(t)G_1(t)\right]v(t)\ dt$$

au seuil

$$\mu = \varkappa + \int_{t_D-T}^{t_D}\left[\tilde{\overline{B}}(t)(G_1(t)-G_o(t))\overline{B}(t)+\tilde{\overline{S}}(t)G_1(t)(\overline{S}(t)+2\overline{B}(t))\right]dt$$

avec décision $\begin{cases} \delta_1 \text{ (choix de } H_1\text{) si } w(t_D) > \mu \\ \delta_o \text{ (choix de } H_o\text{) si } w(t_D) < \mu \end{cases}$

On retrouve alors un modèle de détection par transition dans un espace $\mathcal{Y} = \mathbb{R}$ dont on peut optimaliser les paramètres selon la technique classique de maximalisation d'un rapport signal sur bruit.

5-2-

Ainsi, pour $G_o(t) = G_1(t) = G(t)$, en supposant $\overline{S}(t)$ non identiquement nul, la règle est de comparer la quantité :

$$w(t_D) = \int_{t_D-T}^{t_D}\tilde{\overline{S}}(t)G(t)v(t)\ dt$$

au seuil

$$\mu = \int_{t_D-T}^{t_D} \tilde{\bar{S}}(t)G(t)\left[\frac{1}{2}\,\bar{S}(t) + \bar{B}(t)\right] dt + \frac{\varkappa}{2}$$

Pour G(t) fixée a priori la structure compare à un seuil la sortie du filtre adapté à la référence $\bar{S}(t)$ du signal dans un bruit blanc. Un cas particulier apparaît quand S(t) et E{B(t)} sont complètement connus :

$$G(t) = I \quad , \quad \bar{S}(t)=S(t) \quad , \quad \bar{B}(t) = E\{B(t)\}$$

Pour G(t) non fixée a priori on montre dans [3] que, sous certaines hypothèses, on peut optimaliser le modèle en déterminant le couple $\bar{S}(t)$, G(t), maximalisant un rapport signal sur bruit construit sur $W(t_D)$: on retrouve la comparaison à un seuil de la sortie du filtre adapté au signal dans un bruit blanc.

5-3-

Ainsi, pour $\bar{S}(t) = 0$ et $G(t) = G_o(t) - G_1(t)$ définie positive ou nulle mais non identiquement nulle, la règle est de comparer la quantité :

$$w(t_D) = \int_{t_D-T}^{t_D} \left[\tilde{v}(t)-\tilde{\bar{B}}(t)\right]G(t)\left[v(t)-\bar{B}(t)\right] dt$$

au seuil $\mu = \varkappa$, conduisant à une structure quadratique. Un cas particulier apparaît quand S(t) est aléatoire et que E{B(t)} = E{S(t)} = 0 en posant :

$$\bar{S}(t) = \bar{B}(t) = 0$$

Pour G(t) non fixée a priori on montre dans [3] que sous certaines hypothèses on peut optimaliser le modèle en déterminant G(t) maximalisant un rapport signal sur bruit construit sur $W(t_D)$.

5-4-

Ainsi, pour une hypothèse H_1 composée, en estimant le paramètre θ dont dépend le signal, selon les moindres carrés pondérés par $G_1(t)$:

$$\bar{S}(t) = \bar{S}(t,\hat{\theta}_{MC}) \ ,$$

la règle de détection est de comparer la quantité :

$$w(t_D) = \int_{t_D-T}^{t_D} \left[\tilde{v}(t)-\tilde{\bar{B}}(t)\right]\left[G_o(t)-G_1(t)\right]\left[v(t)-\bar{B}(t)\right] dt$$

$$+ \sup_{\theta} \int_{t_D-T}^{t_D} 2\,\tilde{\bar{S}}(t,\theta)G_1(t)\left[v(t)-\bar{B}(t)-\frac{1}{2}\,\bar{S}(t,\theta)\right]dt$$

au seuil $\mu = \varkappa$.

On détermine de cette manière, par exemple, la structure de systèmes pour la détection, dans une observation v(t), d'un signal n_e-vectoriel certain S(t) dépendant de paramètres θ inconnus, perturbé par un bruit n_e-vectoriel additif B(t) pour lequel l'information est réduite à la connaissance du premier moment. On pose :

$$G_o(t) = G_1(t) = I \ , \ \bar{B}(t) = E\{B(t)\} \ , \ \bar{S}(t) = S(t;\hat{\theta}_{MC})$$

Il en résulte :

• Signal d'amplitude inconnu : S(t) = A Z(t) où Z(t) est déterministe connue et la (n_e,n_e)-matrice diagonale A inconnue. On obtient le test (figure 3)

$$w(t_D) = \sum_{j=1}^{n_e} \frac{\left[\int_{t_D-T}^{t_D} Z_j(t)\left[v_j(t)-E\{B_j(t)\}\right] dt\right]^2}{\int_{t_D-T}^{t_D} Z_j^2(t)\, dt} \ \underset{\delta_o}{\overset{\delta_1}{\gtrless}} \ \mu=\varkappa$$

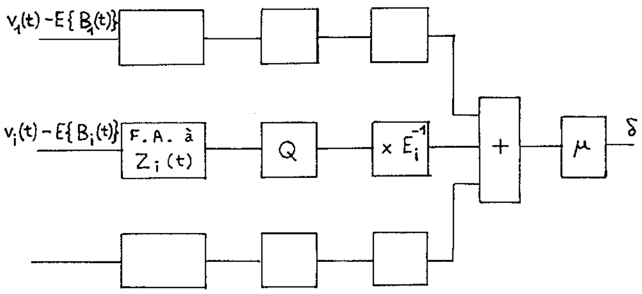

FIGURE 3

• Signal de direction inconnue : $S_j(t,\theta)=Z_j(t-\theta_j)$ où les n_e composantes Z_j sont connues et les n_e retards θ_j sont inconnus . On suppose $|\theta_j| \leqslant \theta_o$ pour tout j, les Z_j de support toujours contenu dans $[t_D-T,t_D]$ et E{B(t)} lentement variable sur θ_o . On obtient le test (figure 4)

$$w(t_D) = \sup_{\{\theta_j\}} \sum_{j=1}^{n_e} \int_{t_D-T}^{t_D} Z_j(t-\theta_j)v_j(t)\, dt \ \underset{\delta_o}{\overset{\delta_1}{\gtrless}}$$

$$\mu = \frac{\varkappa}{2} + \frac{1}{2}\int_{t_D-T}^{t_D} \tilde{Z}(t)Z(t)\, dt + \int_{t_D-T}^{t_D} \tilde{Z}(t)\, E\{B(t)\}\, dt$$

APPLICATION EN DETECTION D'UN CRITERE DE MINIMALISATION DE DISTANCE D'OBSERVATION

• Signal à bande étroite à phases et fréquences inconnues et quelconques : $S_j(t)=A_j(t)\cos(2\pi\nu_j+\varphi_j)$ où les $A_j(t)$ sont connues et les phases φ_j et fréquences ν_j sont inconnues et quelconques. On suppose T grand devant les périodes, et $E\{B(t)\}$ et $A(t)$ lentement variables sur les périodes. On obtient le test (figure 5)

$$w(t_D) = \sum_{j=1}^{n_e} \underset{t=t_D}{\mathcal{E}nv} \int_{t_D-T}^{t_D} A_j(t)\cos(2\pi\nu_j t)v_j(t)dt$$

$$\overset{\delta_1}{\underset{\delta_0}{\gtrless}} \mu = \frac{\chi_0}{2} + \frac{1}{4}\sum_{j=1}^{n_e}\int_{t_D-T}^{t_D} A_j^2(t)\,dt$$

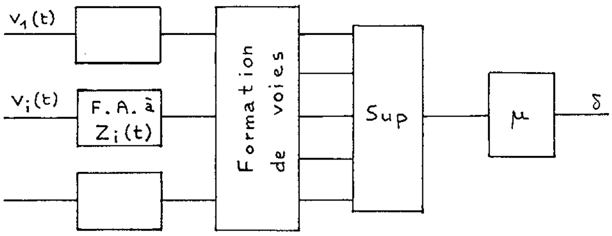

FIGURE 4

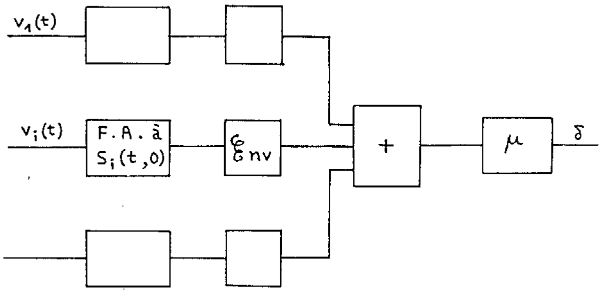

FIGURE 5

BIBLIOGRAPHIE

[1] VAN TREES H.L. : Detection, Estimation and Modulation Theory. Part 1. Wiley, New York, 1968.

[2] ARQUES P.-Y., MACCHI O. VEZZOSI G. : Trois procédures de détection-estimation simultanées d'un signal. Annales des Télécommunications. T. 28, n° 11-12, Nov.-Déc. 1973, p. 459-468.

[3] ARQUES P.-Y. : Décision par minimalisation de distance d'observation. A paraître .