

Détection statistique optimale dans un système linéaire en présence de paramètres de nuisance

Mitra FOULADIRAD¹, Igor NIKIFOROV²

^{1,2}Laboratoire de Modélisation et Sécurité des Systèmes, 12 rue Marie Curie - B.P. 2060 - 10010, Troyes, France

Mitra.Fouladirad@utt.fr, Igor.Nikiforov@utt.fr

Résumé – Le but de cette communication est de proposer un outil statistique optimal pour détecter un «signal» dans un système linéaire stochastique (dynamique) en présence des incertitudes (paramètres de nuisance). On suppose que les paramètres de nuisance sont inconnus et non aléatoires, ce qui signifie qu'en pratique ces derniers peuvent être choisis de façon à ce qu'ils maximisent leur négatif impact sur le système (par exemple, en dissimulant le signal). On illustre les méthodes proposées à l'aide d'un exemple du contrôle de l'intégrité en navigation par satellites.

Abstract – The goal of this paper is to propose an optimal statistical tool to detect a "signal" in a linear stochastic (dynamical) system with uncertainties (nuisance parameters). It is supposed that the nuisance parameters are unknown but non random; practically, this means that the nuisance can be intentionally chosen to maximize its negative impact on the monitored system (for instance, to mask a fault). An example of GNSS integrity monitoring illustrates the proposed method.

1 Introduction

La conception des méthodes de détection dans un système stochastique (dynamique) en présence des incertitudes déterministes (paramètres de nuisance) nécessite la maîtrise des deux exigences essentielles qui sont d'une part une «sensibilité» suffisante aux paramètres informatifs que l'on veut détecter (et localiser) et d'autre part une «insensibilité» suffisante aux paramètres de nuisance. C'est le problème-clé. Différents critères permettent d'évaluer les qualités des algorithmes. Ils sont obtenus par des approches différentes. La communauté «Automatique» s'appuie généralement sur des modèles et des critères déterministes, soit dans le domaine temporel, soit dans le domaine fréquentiel. Elle a surtout centré ses efforts sur la génération d'«indicateurs» de détection, et moins abordé l'étude de la décision qui doit suivre cette étape. Les travaux dans ce domaine proposent une approche fondée sur des considérations géométriques qui conduisent à formuler le compromis «sensibilité/robustesse» en termes d'optimisation d'indicateurs de détection. Se basant sur une approche statistique, la communauté «Traitement du Signal» pose le problème directement au niveau de la procédure de décision en présence de paramètres de nuisance mais les aspects liés à la prise en compte de la structure des modèles et à la géométrie des sous-espaces associés sont moins développés. Les premiers résultats de comparaison entre d'une part les démarches de conception d'un indicateur de détection «optimal» au sens géométrique et d'autre part les démarches de conception d'une décision statistique optimale sont disponibles [1, 3, 4, 6].

L'objectif de cette communication est de continuer cet axe de recherche et de présenter les premiers éléments de la théorie unifiée de détection dans un système stochastique (dynamique) avec redondance en présence des paramètres de nuisance. On suppose que les paramètres de nuisance sont inconnus et non aléatoires, ce qui signifie qu'en pratique ces derniers peuvent

être choisis de façon à ce qu'ils maximisent leur impact négatif sur le système à contrôler (par exemple, en masquant le signal à détecter). On illustre les méthodes proposées avec un exemple du contrôle de l'intégrité du système GNSS à base de station au sol ou autonome (à bord d'un avion). Le champ potentiel d'applications de cette théorie est plus vaste que la problématique de surveillance de systèmes industriels ou le contrôle d'intégrité de systèmes de navigation. Les retombées de cette théorie doivent également permettre d'élaborer les méthodes optimales de détection multi-utilisateur dans un système de communication du type CDMA, par exemple, pour détecter l'apparition d'un nouvel utilisateur dans le système CDMA.

2 Position du problème

L'objet de cette communication est un test statistique (ou règle de décision) de choix entre 2 hypothèses : l'hypothèse de base \mathcal{H}_0 et la contre-hypothèse \mathcal{H}_1 , c'est-à-dire une application surjective de l'espace des échantillons $Y = (Y_1, \dots, Y_n)^T$ sur l'ensemble des hypothèses envisagées $\delta : Y \mapsto \{\mathcal{H}_0, \mathcal{H}_1\}$. On suppose que $Y \sim P_\theta$, où $\mathcal{P} = \{P_\theta\}_{\theta \in \Theta}$ est une famille de distribution dépendant d'un paramètre θ . La qualité du test statistique binaire δ pour choisir entre $\mathcal{H}_0 = \{Y \sim P_\theta | \theta \in \Theta_0\}$ et $\mathcal{H}_1 = \{Y \sim P_\theta | \theta \in \Theta_1\}$ est définie par la *probabilité de fausse alarme* : $\alpha = \sup_{\theta \in \Theta_0} \Pr_\theta (\delta = \mathcal{H}_1)$ et la *fonction de puissance* : $\beta_\delta(\theta) = \Pr_\theta (\delta = \mathcal{H}_1)$. L'application des tests uniformément les plus puissants (UPP) dans le traitement du signal est relativement limitée car leur existence est conditionnée soit par le rapport de vraisemblance Λ *monotone* par rapport à un paramètre *scalaire* θ et une contre-hypothèse \mathcal{H}_1 unilatérale, soit par la présence d'une famille exponentielle scalaire et une hypothèse de base \mathcal{H}_0 bilatérale [2, 5]. La principale limitation de l'usage des tests optimaux dans le cas vectoriel $\theta \in \mathbb{R}^r$ est la «richesse» des domaines Θ_0 et/ou Θ_1 . Pour résoudre le pro-

blème de test des hypothèses composées avec des paramètres de nuisance nous utilisons la théorie développée par Wald dans son article [9] et on l'adapte au cas où le paramètre informatif θ et la nuisance X appartiennent aux différents sous-espace de l'espace d'observations.

3 Modèle linéaire simple

Pour commencer nous supposons le modèle de mesures suivant (sans paramètre de nuisance) :

$$Y = M\theta + \xi, \quad (1)$$

où $Y \in \mathbb{R}^n$, $\theta \in \mathbb{R}^r$ est le paramètre informatif, M une matrice de rang plein de taille $(n \times r)$ avec $r < n$ et $\xi \sim \mathcal{N}(0, \sigma^2 I_n)$ avec la variance $\sigma^2 > 0$ connue. L'objectif est de choisir entre les hypothèses $\mathcal{H}_0 : \{\theta = 0\}$ et $\mathcal{H}_1 : \{\theta \neq 0\}$ en s'inspirant de la méthode de Wald. Le résultat principal est présenté dans le théorème suivant :

Théorème 1. Soit

$$\delta^*(Y) = \begin{cases} \mathcal{H}_0 & \text{si } \Lambda(Y) = \hat{\theta}^T \mathcal{F}_{\hat{\theta}} \hat{\theta} < h(\alpha) \\ \mathcal{H}_1 & \text{si } \Lambda(Y) = \hat{\theta}^T \mathcal{F}_{\hat{\theta}} \hat{\theta} \geq h(\alpha) \end{cases} \quad (2)$$

où $\hat{\theta} = (M^T M)^{-1} M^T Y$ est l'estimateur du maximum de vraisemblance de θ et $\mathcal{F}_{\hat{\theta}} = \frac{1}{\sigma^2} M^T M$ la matrice de Fisher. Le test $\delta^*(Y)$ est UPP dans la classe des tests de niveau α ayant une fonction puissance $\beta_\delta(\theta)$ constante sur chaque surface $S_c \in \mathcal{S}$, avec la famille de surfaces (ellipsoïdes) suivante :

$$\mathcal{S} = \left\{ S_c : \theta^T \mathcal{F}_{\hat{\theta}} \theta = \frac{1}{\sigma^2} \theta^T M^T M \theta = c^2 \right\}$$

Le test $\delta^*(Y)$ coïncide avec le test du rapport de vraisemblance généralisé (RVG). La statistique $\Lambda(Y) = \frac{1}{\sigma^2} Y^T M (M^T M)^{-1} M^T Y$ est distribuée selon une loi du χ_r^2 centrée de r degrés de liberté sous l'hypothèse \mathcal{H}_0 et non centrée $\chi_{r,\lambda}^2$ sous \mathcal{H}_1 avec le paramètre de non-centralité $\lambda = c^2 = \frac{1}{\sigma^2} \theta^T M^T M \theta$. La fonction de puissance est

$$\beta_{\delta^*}(c^2) = \Pr_{c^2}(\delta^* = \mathcal{H}_1) = \Pr(\chi_{r,c^2}^2 \geq h).$$

Le seuil $h = h(\alpha)$ est choisi comme solution de l'équation suivante : $\alpha = \Pr_0(\delta^* = \mathcal{H}_1) = \Pr(\chi_r^2 \geq h)$ où α est la probabilité de fausse alarme. La fonction de puissance typique et les surfaces de puissances constantes sont présentées sur la figure 1.

4 Modèle linéaire avec des paramètres de nuisance

Supposons le modèle linéaire suivant :

$$Y = HX + M\theta + \xi, \quad (3)$$

où $Y \in \mathbb{R}^n$, $\theta \in \mathbb{R}^r$ est le vecteur informatif, M est une matrice de rang plein de taille $(n \times r)$ avec $r < n$, $X \in \mathbb{R}^m$ est le vecteur d'état inconnu et non aléatoire (paramètre de nuisance), H est une matrice de rang plein de taille $(n \times m)$ avec $m < n$, et $\xi \sim \mathcal{N}(0, \sigma^2 I_n)$ avec la variance $\sigma^2 > 0$ connue. Le but est de choisir entre les hypothèses suivantes :

$$\begin{aligned} \mathcal{H}_0 & : \{Y \sim \mathcal{N}(HX, \sigma^2 I_n), X \in \mathbb{R}^m\} \text{ et} \\ \mathcal{H}_1 & : \{Y \sim \mathcal{N}(HX + M\theta, \sigma^2 I_n), \theta \neq 0, X \in \mathbb{R}^m\}. \end{aligned} \quad (4)$$

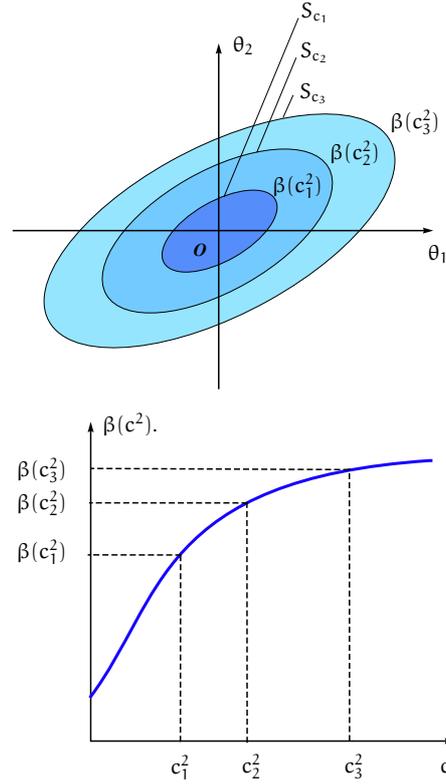


FIG. 1 – La famille des ellipsoïdes et la fonction de puissance $\beta(c^2)$.

La distribution de $Y \sim \mathcal{N}(HX + M\theta, \sigma^2 I_n)$ est invariante par rapport au groupe de translations $G = \{g : Y \mapsto Y + HC\}$, $C \in \mathbb{R}^m$. Le groupe $\bar{G} = \{\bar{g} : \mathbb{E}(Y) \mapsto \mathbb{E}(Y) + HB\}$ préserve les ensembles de paramètres :

$$\begin{aligned} \Omega_0 & = \{\mathbb{E}(Y) = HX, X \in \mathbb{R}^m\} \text{ et} \\ \Omega_1 & = \{\mathbb{E}(Y) = HX + M\theta, X \in \mathbb{R}^m, \theta \neq 0\}, \end{aligned}$$

i.e. $\bar{g}\Omega_0 = \Omega_0$ et $\bar{g}\Omega_1 = \Omega_1$. Donc le problème (4) est invariant. Par conséquent tous les tests invariants optimaux sont basés sur l'invariant maximal. Soit $R(H)$ l'espace engendré par les colonnes de la matrice H . Le vecteur de parité $Z = WY$ est la projection de Y sur $R(H)^\perp$ (le complément orthogonal de $R(H)$ dit «l'espace de parité»). Ce principe est présenté sur la figure 2 dans le cas où $Y \in \mathbb{R}^3$ et $X \in \mathbb{R}$. La matrice de projection $W = (w_1, \dots, w_{n-m})^T$ de taille $(n-m) \times n$ est composée des vecteurs propres w_1, \dots, w_{n-m} correspondant à la valeur propre 1 de la matrice de projection $P_H = I_n - H(H^T H)^{-1} H^T$.

$$WH = 0, \quad W^T W = P_H, \quad WW^T = I_{n-m}.$$

Le vecteur $Z = WY$ est invariant maximal relativement au groupe de translations $G = \{g : Y \mapsto Y + HX\}$. Donc tous les tests invariants doivent dépendre de Y seulement via le vecteur $Z = WY$ [5].

Théorème 2. Soit la matrice WM de rang plein de taille $((n-m) \times r)$, $r < n-m$. Le test

$$\delta_H^*(Z) = \begin{cases} \mathcal{H}_0 & \text{si } \Lambda(Z) < h(\alpha) \\ \mathcal{H}_1 & \text{si } \Lambda(Z) \geq h(\alpha) \end{cases} \quad (5)$$

où $\Lambda(Z) = \frac{1}{\sigma^2} Z^T WM ((WM)^T WM)^{-1} M^T W^T Z$, est UPP

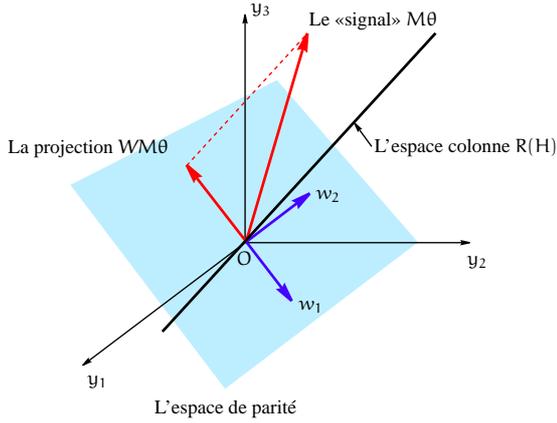


FIG. 2 – L'espace engendré par les colonnes de H et son complément orthogonale («espace de parité»).

invariant avec la puissance $\beta_\delta(\theta)$ constante sur chaque surface $S_c \in \mathcal{S}$ de la famille de surfaces :

$$\mathcal{S} = \left\{ S_c : \frac{1}{\sigma^2} \theta^T M^T W^T W M \theta = c^2 \right\}.$$

Le test $\delta_{\mathcal{H}}^*(Z)$ coïncide avec le test RVG. Les propriétés du test RVG dans le cadre du modèle (3) ont été étudiées dans l'article [7]. Pour démontrer l'optimalité du test RVG les auteurs de l'article [7] utilisent le fait que la distribution de la statistique du test RVG (loi du χ^2) est une fonction monotone du paramètre de non-centralité λ . C'est-à-dire les auteurs [7] ont réduit le problème de détection du vecteur (paramètre informatif) θ au problème de détection du paramètre scalaire $\lambda = \frac{1}{\sigma^2} \theta^T M^T W^T W M \theta$ sans garantir que cette réduction laisse le sens d'optimalité du problème inchangé. Cela nous laisse à penser que la réduction du vecteur θ au paramètre λ est heuristique. L'avantage de notre approche par rapport au résultat de l'article [7] est la démonstration d'optimalité directement dans l'espace du paramètre informatif θ .

5 Système stochastique dynamique avec des paramètres de nuisance

En utilisant la théorie développée, nous présentons ici l'analyse de l'impact du paramètre de nuisance représenté par une équation d'état *déterministe*. On considère le modèle d'état suivant

$$X_k = F X_{k-1} + B U_k, \quad (6)$$

$$Y_k = H X_k + M \theta_k + \xi_k, \quad (7)$$

où $Y_k \in \mathbb{R}^n$ est le vecteur des mesures, $X_k \in \mathbb{R}^m$ est le vecteur d'état, $U_k \in \mathbb{R}^p$ représente le vecteur des entrées inconnues (paramètre de nuisance), $\theta_k \in \mathbb{R}^r$ est le paramètre informatif à l'instant k , $\xi_k \sim \mathcal{N}(0, \sigma^2 I_n)$, la variance $\sigma^2 > 0$ est connue et la matrice de mesures H est de rang plein et de taille $(n \times m)$ avec $m < n$. Considérons un échantillon $\mathcal{Y} = (Y_1, \dots, Y_N)$. Il existe deux méthodes pour traiter le problème de détection des vecteurs $\theta_1, \dots, \theta_N$ en présence des paramètres de nuisance $X_1 = X_1(U_1, X_0), \dots, X_N = X_N(U_1, \dots, U_N, X_0)$. La première consiste à ignorer la présence de l'équation d'état et la seconde – à utiliser l'équation de mesure et l'équation d'état ensemble.

5.1 Ignorer la présence de l'équation d'état

Cette méthode est basée sur l'utilisation seulement l'équation de mesure (7). On considère le modèle

$$y = \mathcal{H}X + M\theta_{1,N} + \xi_{1,N}.$$

où $X = (X_1^T, \dots, X_N^T)^T$, $\theta_{1,N} = (\theta_1^T, \dots, \theta_N^T)^T$, $\xi_{1,N} = (\xi_1^T, \dots, \xi_N^T)^T$ et les matrices M et H ont respectivement M et H en diagonale. Le test $\delta_{\mathcal{H}}^*$ basé sur la matrice de projection W telle que $W\mathcal{H} = 0$, $W^T W = \mathcal{P}_{\mathcal{H}}$ et $W W^T = I_{Nn-Nm}$ est UPP invariant dans la classe des tests de niveau α ayant une fonction puissance $\beta_\delta(\theta)$ constante sur chaque surface $S_c \in \mathcal{S}_{\mathcal{H}}$, avec la famille de surfaces :

$$\mathcal{S}_{\mathcal{H}} = \left\{ S_c : \frac{1}{\sigma^2} \| W M \theta_{1,N} \|_2^2 = c^2 \right\}.$$

5.2 Utiliser l'équation de mesure et l'équation d'état ensemble

Dans le cadre de cette méthode, le détecteur est basé sur les deux équations (de mesure (7) et d'état (6)). On obtient le modèle suivant

$$\begin{aligned} y &= \mathcal{U}X_0 + \mathcal{V}U_{1,N} + M\theta_{1,N} + \xi_{1,N} \\ &= (\mathcal{U} \quad \mathcal{V}) \begin{pmatrix} X_0 \\ U_{1,N} \end{pmatrix} + M\theta_{1,N} + \xi_{1,N}. \end{aligned}$$

où $U_{1,N} = (U_1^T, \dots, U_N^T)^T$ et les matrices U et V sont des fonctions des matrices H, F et B. Le test $\delta_{(\mathcal{U} \quad \mathcal{V})}^*$ basé sur la matrice de projection W_1 telle que $W_1(\mathcal{U} \quad \mathcal{V}) = 0$, $W_1 W_1^T = I_{Nn-\text{rang}(\mathcal{U} \quad \mathcal{V})}$ et $W_1^T W_1 = \mathcal{P}_{(\mathcal{U} \quad \mathcal{V})}$ est UPP invariant dans la classe des tests de niveau α ayant une fonction puissance $\beta_\delta(\theta)$ constante sur chaque surface $S_c \in \mathcal{S}_{(\mathcal{U} \quad \mathcal{V})}$, avec la famille de surfaces :

$$\mathcal{S}_{(\mathcal{U} \quad \mathcal{V})} = \left\{ S_c : \frac{1}{\sigma^2} \| W_1 M \theta_{1,N} \|_2^2 = c^2 \right\}.$$

La comparaison de ces deux méthodes est présentée dans le théorème suivant :

Théorème 3. Soit $\Pr_0(\delta_{\mathcal{H}}^* = \mathcal{H}_1) = \Pr_0(\delta_{(\mathcal{U} \quad \mathcal{V})}^* = \mathcal{H}_1) = \alpha$, ($0 < \alpha < 1$). Alors $\forall \theta_{1,N} \neq 0$, $\beta_{\delta_{\mathcal{H}}^*}(\theta_{1,N}) \leq \beta_{\delta_{(\mathcal{U} \quad \mathcal{V})}^*}(\theta_{1,N})$.

Remarques : Ce théorème montre l'importance d'information a priori apportée par l'équation d'état (6) sur les paramètres de nuisance «dynamiques». Dans la littérature sur redondance analytique l'importance de l'équation d'état a été soulignée (voir [8]), mais jamais démontrée formellement. Le théorème 3 prouve ce résultat. Néanmoins, ce n'est pas certain a priori qu'on puisse toujours améliorer la puissance du test UPP invariant avec puissance constante basé seulement sur l'équation de mesure (7) en utilisant les deux équations (7) et (6).

6 Application : contrôle de l'intégrité du GNSS

Un système de navigation est un dispositif ou une combinaison de dispositifs embarqués dont le but fondamental est de fournir des informations de navigation par rapport à un point

géographique sélectionné. Les quatre paramètres de base employés pour définir les critères de performance des systèmes de navigation sont : la précision, l'intégrité, la continuité et la disponibilité. Nous allons discuter ici l'intégrité d'un système de navigation. La définition de ce paramètre est la suivante : le degré de confiance que l'on peut placer dans l'exactitude des informations fournies par le système, ainsi que la capacité à avertir l'utilisateur d'un dysfonctionnement du système (lorsqu'une partie ou la totalité de ses composants fournissent des renseignements erronés) dans un délai raisonnable pour signaler que le système ne devrait pas être utilisé pour l'opération (ou la phase de vol) prévue. Pour les applications aéronautiques, le contrôle de l'intégrité des signaux GNSS est très important, en particulier pour les phases de vol critiques au niveau de sécurité (atterrissage, décollage, ...). Le principe du contrôle d'intégrité repose sur l'utilisation de la redondance des informations fournies par les mesures de distances. En utilisant la théorie développée, nous présentons ici deux méthodes de contrôle d'intégrité (à base de station au sol et à bord d'un avion) et l'analyse de l'impact d'un modèle d'horloge représenté par une équation d'état pour illustrer le théorème 3.

6.1 Contrôle de l'intégrité du GNSS à base de station au sol

Supposons que l'on dispose de n satellites, placés à des positions connues $X_i = (x_i, y_i, z_i)^T$, $i = 1, \dots, n$, visibles par un récepteur (station au sol) dont la position 3D (connue) est $X_r = (x_r, y_r, z_r)^T$. L'équation de mesure est donc

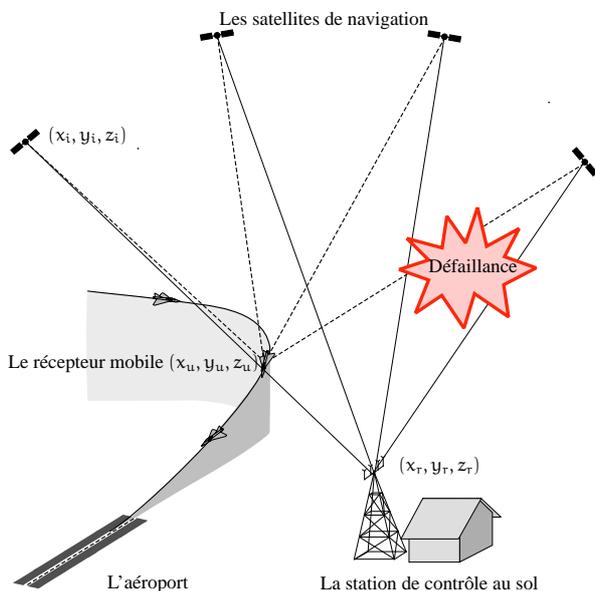


FIG. 3 – Contrôle d'intégrité du GNSS à base de station au sol.

$$y_i = r_i - d_i = c \rho + \xi_i, \quad \xi_i \sim \mathcal{N}(0, \sigma^2), \quad i = 1, \dots, n,$$

où r_i est la pseudo-distance (distance mesurée) entre le satellite i et le récepteur, $d_i = \|X_i - X_r\|_2$ est la vraie distance entre le satellite i et le récepteur, $c \simeq 2.9979 \cdot 10^8$ m/s est la vitesse de la lumière, ρ est le biais inconnu et non aléatoire d'horloge du récepteur par rapport au temps GNSS, ξ_i est le bruit de pseudo-distance. En ajoutant le vecteur des anomalies

de propagation θ , on peut réécrire l'équation de mesure sous la forme suivante :

$$Y = C\rho + \xi + \theta \quad (8)$$

où $C = (c, \dots, c)^T$, $Y = (y_1, \dots, y_n)^T$, $\xi = (\xi_1, \dots, \xi_n)^T$. Le problème du contrôle des signaux GNSS à base de station au sol consiste à détecter la présence du vecteur θ qui peut être considéré comme défaillance de positionnement.

6.2 Contrôle autonome de l'intégrité du GNSS

On considère maintenant le contrôle autonome (à bord d'un avion) de l'intégrité du GNSS. Supposons que la position 3D $X_u = (x_u, y_u, z_u)^T$ d'un récepteur mobile est inconnue. Le modèle de mesures linéarisé est

$$Y = H(X - X_0) + \xi + \theta, \quad (9)$$

où $Y = (y_1, \dots, y_n)^T$, H est la matrice Jacobienne de dimension $n \times 4$ composée des cosinus directeurs récepteur-satellite (trois premières colonnes) et des c (quatrième colonne), $X = (X_u^T, \rho)^T$ représente le vecteur d'état inconnu, X_0 est le point de travail et $\xi \sim \mathcal{N}(0, \sigma^2 I_n)$ est le vecteur de bruit de pseudo-distance.

Le problème du contrôle d'intégrité autonome (à bord d'un avion) des mesures GNSS consiste à détecter une anomalie de propagation des signaux émis par les satellites, c'est-à-dire la détection de la présence d'un vecteur θ qui peut être considéré comme défaillance de positionnement.

Références

- [1] M. Basseville and I. V. Nikiforov. Fault isolation for diagnosis : Nuisance rejection and multiple hypothesis testing. *Annual Reviews in Control*, 26(2) :189–202, 2002.
- [2] A. A. Borovkov. *Mathematical Statistics*. Gordon and Breach Sciences Publishers, Amsterdam, 1998.
- [3] H. Lacresse, A. Grall, I. Nikiforov. Passerelles entre les approches statistiques et géométriques de la détection. In *GRETSI'01, Toulouse, France, CD ROM*, 2001.
- [4] I. Nikiforov (Resp.). Rapport du Projet 2 de S3 : Sécurité, Surveillance, Supervision : Etude des complémentarités entre redondance analytique et décision statistique pour la conception d'algorithmes optimaux de détection et de localisation des défaillances. In *Journées Nationales d'Automatique, Autrans, CD Rom*, 31 janvier - 02 février 2001.
- [5] E. L. Lehmann. *Testing Statistical Hypotheses*. Chapman and Hall, 1986.
- [6] I. Nikiforov. Eléments de théorie de la décision statistique II : compléments, Chap. 2. In *Décision et reconnaissance des formes en signal*, pages 43–84. Hermes Science Publications (Traité IC2), Edited by R. Lengellé, 2002.
- [7] L.L. Scharf and B. Friedlander. Matched subspace detectors. *IEEE Trans. Sign. Proc.*, 42(8) :2146–2157, 1994.
- [8] M. Staroswiecki. Redondance analytique. In *Automatique et statistiques pour le diagnostic*, pages 43–68. Hermes Science Europe, 2001.
- [9] A. Wald. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Amer. Math. Soc.*, 54 :426–482, 1943.