



HAL
open science

Actes du groupe de travail en biostatistiques NANCY septembre 2005-juin 2006

Pierre P. Vallois, Jean-Marie Monnez

► **To cite this version:**

Pierre P. Vallois, Jean-Marie Monnez. Actes du groupe de travail en biostatistiques NANCY septembre 2005-juin 2006. 2007. hal-00135156

HAL Id: hal-00135156

<https://hal.science/hal-00135156>

Preprint submitted on 30 Jun 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Introduction

Dans les Actes de l'année précédente était évoqué le problème initial qui a motivé la constitution de notre groupe de travail. A l'aide d'un spectromètre de masse, on détermine à partir d'un échantillon de sang d'un individu les quantités de protéines de différents types présentes. Le problème est alors de voir si l'on peut distinguer par cette analyse (les statisticiens disent "discriminer") un individu sain d'un individu atteint d'un certain type de cancer, de déterminer des protéines qui sont des "biomarqueurs" de ce cancer (il s'agit de déterminer des caractères discriminants), puis, pour un individu dont on ne sait pas s'il est sain ou cancéreux, de le classer dans l'une de ces catégories à partir de l'observation des biomarqueurs (c'est le problème du classement, encore appelé classification supervisée). L'intérêt statistique de l'étude est que l'on peut observer dans un spectre 15000 quantités de protéines différentes, alors que le nombre d'individus observés n'est que de deux ou trois centaines: on a un tableau de données avec peu d'individus et un très grand nombre de caractères (c'est le "fléau de la dimension").

Comme tout grand tableau de données statistiques, il devra être "toileté" pour éliminer éventuellement des individus aberrants ou des caractères fortement corrélés avec d'autres et qui apportent donc peu d'information nouvelle. Il est intéressant de "voir" ce tableau, de voir les proximités et les différences entre individus. Tout ceci peut être réalisé en effectuant une analyse en composantes principales, qui permet également de remplacer un très grand nombre de caractères par un nombre réduit de facteurs, combinaisons linéaires des caractères, non corrélés, qui peuvent alors être utilisés dans une autre étude statistique, comme une analyse discriminante, à la place des caractères. Cependant, l'inconvénient que voit le praticien est que ces facteurs n'ont pas une signification physique précise (on peut cependant les interpréter par leurs corrélations avec les caractères), mais sont des synthèses de l'ensemble des caractères; le statisticien quant à lui voit l'avantage d'avoir un nombre réduit de facteurs à la place d'un grand nombre de caractères.

Les contributions placées en tête de ces Actes (les Chapitres de 1 à 6) correspondent à des cours et leur intérêt est de définir précisément, du point de vue mathématique, différents outils statistiques susceptibles d'être utilisés dans nos travaux de recherche. Pour les raisons évoquées plus haut, nous nous sommes concentrés sur les problèmes de discrimination. Bien qu'il s'agisse de notions déjà connues, il nous a paru utile de rassembler dans un même volume ces éléments de statistique parfois éparpillés dans la littérature. De

plus ce document pourrait servir de référence pour un plus large public que les mathématiciens; nous pensons en particulier aux biologistes, bio-informaticiens et automaticiens qui participent à notre groupe.

Détaillons brièvement l'organisation de la première partie de ce document. Dans le Chapitre 1, est présentée une méthode de sélection de caractères discriminants (dans les Actes de l'année dernière figurait l'analyse de la variance à un facteur qui est une de ces méthodes). Dans le Chapitre 2, est étudiée la construction d'une fonction discriminante entre deux classes par la méthode des SVM ("support vector machines" ou "séparateurs à vaste marge"), et dans le Chapitre 3 une méthode de sélection de caractères discriminants basée sur cette approche. Dans les Chapitres 4 et 5, on présente deux articles qui ont trait à la détermination des biomarqueurs protéiques du cancer des ovaires. Dans le Chapitre 6 est présentée l'analyse en composantes principales.

Notre groupe a cherché à s'ouvrir vers l'extérieur en invitant en particulier des conférenciers non-mathématiciens. Afin de susciter des collaborations, ces derniers nous ont exposé leurs travaux en soulignant l'utilisation de probabilités et de statistique. Certains orateurs ont accepté de rédiger leurs interventions. Qu'ils en soient remerciés.

Nous avons regroupé ces textes dans les Chapitres numérotés de 7 à 10. On trouvera dans le Chapitre 7 un court texte correspondant à un exposé de Mme Luposi du Centre Alexis Vautrin. La conférence donnée par T. Bastogne du CRAN figure dans le Chapitre 8 et le sujet est l'analyse pharmacocinétique de la vitesse d'incorporation intra-cellulaire de molécules photoactivables. Cet exposé a été suivi d'un travail en commun avec S. Wantz et P. Vallois et a fait l'objet d'une publication en cours de soumission à l'heure actuelle. Dans le Chapitre 9, V. Louis-Dorr et R. Ranta du CRAN présentent une modélisation de l'analyse de signaux cérébraux de patients épileptiques à l'aide de la théorie probabiliste du signal. Pour terminer, le Chapitre 10 est la version écrite de l'exposé de W. Blondel du Centre de Recherche en Automatique de Nancy (CRAN) dont l'objet est la spectroscopie tissulaire pour le diagnostic in vivo : simulations statistiques des interactions lumière-tissu en diffusion élastique et autofluorescence pour identifier les propriétés optiques des tissus. J.S. Giet et S. Tindel ont entrepris une collaboration scientifique avec W. Blondel.

Il ne nous pas été possible d'avoir une rédaction de tous les exposés effectués au groupe de travail pendant l'année 2006. Toutefois certaines conférences de 2006 seront reprises et complétées, puis seront publiées dans les Actes de l'année prochaine.

Nancy le 8 février 2007

Jean-Marie Monnez et Pierre Vallois.

TABLE DES MATIERES.

1. Introduction.....	1
2. Méthodes de sélection de caractères discriminants. <i>Jean-Marie MONNEZ</i>	4
3. Eléments d'étude des Séparateurs à Vaste Marge (S.V.M.). <i>Jean-Marie MONNEZ, Sophie Wantz-Mézières</i>	17
4. Sélection de composantes pour les SVM. <i>Olivier Collignon</i>	25
5. Méthodes statistiques de "Three Biomarkers Identified From Serum Proteomic Analysis for the Detection of Early Stage Ovarian Cancer." de Z.Zhang. <i>Olivier Collignon</i>	32
6. Méthodes statistiques de "Use of proteomic pattern in serum to identify ovarian cancer." de E.F. Petricoin et "Reproducibility of SELDI-TOF protein patterns in serum..." de K.A. Baggerly... <i>Olivier Collignon</i>	37
7. L'analyse en composantes principales. <i>Jean-Marie MONNEZ</i>	43
8. Résumé de la présentation de Elisabeth Luporsi. <i>Elisabeth Luporsi</i>	68
9. Parameter Estimation of Pharmacokinetics Models in the presence of Timing Noise. <i>Thierry Bastogne, Sophie Mézières-Wantz, et al</i>	69
10. Traitement de signaux physiologiques. <i>M. Caparos, V. Louis-Dorr et al</i>	78
11. Photodiagnostic tissulaire in vivo : modélisation des interactions lumière-tissus en spectroscopie d'autofluorescence et de diffusion élastique pour identifier les propriétés optiques de tissus biologiques <i>Walter Blondel, Emilie Péry et al</i>	95

Méthodes de sélection de caractères discriminants.

Jean-Marie MONNEZ

1 Formulation

1.1 Données et but

Soit :

- $I = \{1, \dots, n\}$ un ensemble de n individus, appelé ensemble d'apprentissage (on a donc une méthode supervisée);
- x^1, \dots, x^p p caractères quantitatifs mesurés sur les individus de I ;
- y un caractère qualitatif à q modalités exclusives, observé sur les individus de I , qui induit une partition de I en classes I_1, \dots, I_q .

Exemple. Un ensemble d'individus est divisé en deux classes:

I_1 : classe des cancéreux;

I_2 : classe des non-cancéreux;

x^1, \dots, x^p sont les hauteurs des pics d'un spectre protéomique. ■

Le but est de sélectionner parmi x^1, \dots, x^p des caractères qui discriminent les classes I_1, \dots, I_q .

Cette sélection s'effectue en deux étapes :

- 1) l'élimination de caractères non discriminants est étudiée dans le paragraphe 2;
- 2) la sélection de caractères discriminants parmi ceux retenus est étudiée dans les paragraphes 3 (sélection d'un sous-ensemble de caractères), 4 (sélection pas à pas) et 5 (discrimination logistique binaire; on ne traite pas la discrimination logistique polytomique).

1.2 Traitements préliminaires

Les caractères x^1, x^2, \dots, x^p peuvent être des facteurs issus d'une analyse factorielle, donc des combinaisons linéaires de caractères quantitatifs ou d'indicatrices de modalités. Ce sera en particulier le cas si les caractères initiaux sont qualitatifs.

Il est logique de se demander d'abord si l'ensemble des caractères x^1, \dots, x^p permet de bien discriminer les classes I_1, \dots, I_q . On pourrait avoir une représentation visuelle de la discrimination en faisant une analyse factorielle discriminante (AFD). Les facteurs d'une AFD sont vecteurs propres de $T^{-1}B$ (T : matrice d'inertie totale ; B : matrice d'inertie inter). Si le nombre de caractères est supérieur au nombre d'individus, la matrice T n'est pas de plein rang : l'AFD est donc irréalisable. On verra en appendice comment procéder dans ce cas.

Il est cependant intéressant de visualiser les données. On peut le faire en réalisant une analyse en composantes principales (ACP). Cette analyse permet de repérer les individus aberrants (outliers) à éliminer éventuellement. Elle donne aussi les corrélations entre les caractères initiaux.

On peut aussi partitionner en classes par une méthode de *classification automatique* l'ensemble des caractères, par exemple des classes de caractères fortement corrélés positivement, et effectuer pour chaque classe une AFD.

2 Elimination de caractères

2.1 Tests de comparaison

2.1.1 Comparaison de moyennes

2.1.1.1 Test d'analyse de la variance

Soit un caractère x pris parmi x^1, \dots, x^p .

On note b_1, \dots, b_q les modalités de y .

Soit x_{ij} la $j^{\text{ième}}$ mesure de x lorsque y a la modalité b_i , $j = 1, \dots, n_i$.

On suppose que x_{ij} est la réalisation d'une variable aléatoire réelle (v.a.r.) X_{ij} et qu'il existe des nombres fixes $\mu, \beta_1, \dots, \beta_q$ et une v.a.r. R_{ij} de loi $\mathcal{N}(0, \sigma)$, les R_{ij} étant mutuellement indépendantes, tels que :

$$X_{ij} = \mu + \beta_i + R_{ij}.$$

On effectue alors le test :

$$\begin{cases} H_0 : \beta_1 = \dots = \beta_q \\ H_1 : \overline{H_0}. \end{cases}$$

Si l'on pose $m_i = \mu + \beta_i = E[X_{ij}]$, ceci revient à faire le test :

$$\begin{cases} H_0 : m_1 = \dots = m_q \\ H_1 : \overline{H_0}. \end{cases}$$

Ce test est étudié dans le chapitre d'analyse de la variance à un facteur.

Dans le cas $q = 2$, ce test équivaut au test de comparaison des moyennes de Student (statistique de test : $\frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \sim \mathcal{T}_{n_1 + n_2 - 2}$ sous H_0).

Si l'on ne rejette par l'hypothèse H_0 , on conclut que le caractère x ne discrimine pas de façon significative les classes I_1, \dots, I_q .

Validation a posteriori du modèle

Elle peut être faite en établissant le graphique en fonction de (i, j) des estimations $\frac{x_{ij} - \hat{\mu} - \hat{\beta}_i}{\hat{\sigma}}$ = $\frac{x_{ij} - \bar{x}_i}{\sqrt{\frac{1}{n-q} \sum_i \sum_j (x_{ij} - \bar{x}_i)^2}}$ des résidus réduits $\frac{r_{ij}}{\sigma}$, qui sont distribués suivant

la loi normale centrée réduite, et en observant la répartition des points obtenus.

2.1.1.2 Cas où des hypothèses ne sont pas vérifiées

1) Non égalité des variances (hétéroscédasticité)

On suppose que la variance $\sigma_i^2 = \text{Var}[R_{ij}]$ dépend de la classe I_i . S_i^2 désignant la variance empirique des R_{ij} , $j = 1, \dots, n_i$, on peut estimer σ_i^2 par $\frac{n_i}{n_i - 1} s_i^2 = \hat{\sigma}_i^2$ et remplacer les données x_{ij} par les données $x_{ij}/\hat{\sigma}_i$.

Pour $q = 2$, il existe un test de comparaison des moyennes de deux lois normales d'écart types inconnus et inégaux (tables de Sukhatme; statistique de test : $\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1-1} + \frac{S_2^2}{n_2-1}}}$).

2) Non normalité

a) Transformation du caractère x

On cherche une fonction de x distribuée suivant une loi normale.

b) Cas $q = 2$ avec des tailles d'échantillons n_1 et n_2 suffisamment grandes

On peut utiliser un test de comparaison des moyennes de deux lois quelconques d'écart types inconnus ; il y a deux cas : les écarts-types sont égaux $\left(\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 1} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \right) \sim \mathcal{N}(0, 1)$ sous H_0) ou inégaux $\left(\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1-1} + \frac{S_2^2}{n_2-1}}} \sim \mathcal{N}(0, 1)$ sous H_0).

2.1.2 Comparaison de moyennes multidimensionnelles

Soit un caractère multidimensionnel \underline{x} , constitué de caractères pris parmi x^1, \dots, x^p .

On note b_1, \dots, b_q les modalités de y .

Soit \underline{x}_{ij} la j^{ieme} mesure de \underline{x} lorsque y a la modalité b_i , $j = 1, \dots, n_i$.

On suppose que \underline{x}_{ij} est la réalisation d'un vecteur aléatoire (v.a.) \underline{X}_{ij} et qu'il existe des vecteurs $\underline{\mu}, \underline{\beta}_1, \dots, \underline{\beta}_q$ et un v.a. \underline{R}_{ij} de loi $\mathcal{N}(\underline{0}, \Sigma)$, la matrice de covariance Σ ne

dépendant pas de i et les vecteurs \underline{R}_{ij} étant mutuellement indépendants, tels que :

$$\underline{X}_{ij} = \underline{\mu} + \underline{\beta}_i + \underline{R}_{ij}.$$

On effectue alors le test d'analyse de la variance multidimensionnelle:

$$\begin{cases} H_0 : \underline{\beta}_1 = \dots = \underline{\beta}_q = \underline{0} \\ H_1 : \overline{H}_0. \end{cases}$$

Si l'on pose $\underline{m}_i = \underline{\mu} + \underline{\beta}_i = E[\underline{R}_{ij}]$, ceci revient à faire le test :

$$\begin{cases} H_0 : \underline{m}_1 = \dots = \underline{m}_q \\ H_1 : \overline{H}_0. \end{cases}$$

Ce test est étudié dans la chapitre d'analyse de la variance à un facteur.

2.1.3 Comparaison de lois

On dispose de plusieurs tests de comparaison de lois :

- pour q quelconque, test d'homogénéité du khi-deux, pour des effectifs n_i suffisamment grands;
- dans le cas $q = 2$, test de Kolmogorov-Smirnov ;
- dans le cas $q = 2$, tests de Mann-Whitney et de Wilcoxon ; pour q quelconque, test de Kruskal-Wallis.

Ces tests sont étudiés dans le chapitre sur les tests de comparaison de lois.

Remarque. Une différence entre les variances entraîne une différence entre les lois. Or, l'expérimentateur peut être intéressé en premier lieu par la différence entre les moyennes.

2.1.4 Mesure de la capacité de discrimination d'un caractère

Dans les tests de comparaison au seuil α étudiés, T étant la statistique de test (ou sa valeur absolue), on a une règle de décision de la forme suivante :

"si $T > t_\alpha$, on rejette H_0 ; sinon on ne rejette pas H_0 ".

Soit t_0 la réalisation observée de la statistique T . On définit la *probabilité critique* (ou p -value), notée p :

$$p = P(T > t_0) \quad \text{sous } H_0.$$

Comme $P(T > t_\alpha) = \alpha$ sous H_0 , la règle de décision peut s'écrire ainsi :

"si $p < \alpha$, on rejette H_0 ; sinon on ne rejette pas H_0 ".

Dans le cas où l'on rejette H_0 , la probabilité critique est une mesure de la capacité individuelle de discrimination du caractère (éventuellement multidimensionnel) étudié, sans tenir compte de la présence des autres caractères. On verra plus loin des mesures de la capacité de discrimination en présence d'autres caractères.

2.2 Etude de corrélations

2.2.1 Corrélations avec des facteurs discriminants

On peut réaliser une *AFD* sur un ensemble de caractères (éventuellement des caractères retenus comme discriminants) à condition que le nombre de caractères soit inférieur au nombre d'individus.

Ceci permet de vérifier de visu si les caractères considérés permettent d'effectuer une bonne discrimination. Si les premiers facteurs de l'analyse donnent une bonne discrimination, on peut éliminer des caractères peu corrélés à ces facteurs.

Ainsi, on tient compte dans cette élimination de la présence d'autres caractères.

2.2.2 Corrélations entre caractères discriminants

Parmi un ensemble de caractère discriminants, on peut éliminer des caractères fortement corrélés avec d'autres. Entre deux caractères discriminants, on conserve celui qui l'est le plus.

Remarque 1. On peut aussi remplacer un ensemble de caractères fortement corrélés par le premier facteur d'une *ACP* sur cet ensemble, qui les synthétise tous. Cependant, il peut ne pas avoir une interprétation claire.

Remarque 2. La détermination des corrélations peut se faire dans une *ACP*, qui permet également d'obtenir une représentation graphique des individus.

3 Sélection d'un sous-ensemble de caractères discriminants

On dispose après la première phase d'élimination d'un ensemble de p caractères x^1, \dots, x^p , p désignant dans ce paragraphe le nombre de caractères retenus. Soit r un entier fixé. On va chercher un sous-ensemble de r caractères qui ait la plus grande capacité de discrimination. Ceci nécessite la définition d'une mesure de la capacité de discrimination d'un ensemble de r caractères. Pour déterminer un sous-ensemble optimal, on aura à effectuer C_p^r mesures de cette capacité, ce qui peut entraîner un nombre très important de calculs. On verra dans le paragraphe suivant des méthodes qui réduisent le nombre de calculs.

A l'issue de la sélection d'un sous-ensemble de r caractères, on pourra faire une analyse factorielle discriminante pour vérifier de visu si la discrimination effectuée par les caractères retenus est bonne.

3.1 Critère de l'inertie inter-classes

Chaque individu i de I est représenté par un point de \mathbb{R}^p qui a pour coordonnées les valeurs des caractères x^1, \dots, x^p ; on définit dans \mathbb{R}^p une métrique M .

Si on n'utilise que r caractères pris parmi les p caractères, on se place dans \mathbb{R}^r dans lequel on définit une métrique M_r à laquelle est associée la distance d_{M_r} . On note :

- \underline{x}_r la restriction à \mathbb{R}^r d'un vecteur \underline{x} de \mathbb{R}^p ;
- T_r, B_r, W_r les restrictions à \mathbb{R}^r des matrices d'inertie totale T , inter B , intra W .

On mesure le pouvoir discriminant d'un ensemble de r caractères par l'inertie inter-classes

$$\sum_{k=1}^q P_k d_{M_r}^2(\underline{g}_{k,r}, \underline{g}_r) = \text{Trace}(M_r B_r)$$

($\underline{g}_{k,r}$: restriction de g_k , barycentre de I_k ; P_k poids de I_k ; \underline{g}_r : restriction de \underline{g} , barycentre de I).

On prend en général $M_r = (T_r)^{-1}$.

Règle de sélection. *On retient un ensemble de r caractères qui rend $\text{Trace}(M_r B_r)$ maximale.*

3.2 Critère du pourcentage de bien classés par une règle de classement

1) En utilisant uniquement r caractères, on définit *une règle de classement* :

- géométrique : linéaire, quadratique, par voisinage, par boule ;
- bayésienne ;
- logistique (estimation de $P(\Omega_k | \underline{X} = \underline{x})$);
- par SVM (support vector machine);
- etc...

Ces méthodes sont étudiées dans le chapitre sur les méthodes de classement.

2) On mesure *la qualité de cette règle de classement* par

- la méthode de l'échantillon - test ;
- la validation croisée, le leave-one-out ;
- le bootstrap.

Dans chaque méthode, on calcule un pourcentage de bien classés.

Règle de sélection. *On retient un ensemble de r caractères qui rend le pourcentage de bien classés maximal.*

On détaille ci-après les différentes mesures de la qualité d'une règle de classement.

a) Méthode de l'échantillon-test

On dispose de l'ensemble d'apprentissage I , divisé en classes I_1, \dots, I_q . On suppose que I est un échantillon d'une population Ω , divisé en classes $\Omega_1, \dots, \Omega_q$.

On effectue un sondage par stratification dans I : on tire au hasard de chaque classe I_k des individus, en nombre proportionnel à $\text{card } \Omega_k$, ou à défaut à $\text{card } I_k$; ces individus constituent l'échantillon-test. On peut prendre jusqu'à 30 % des individus de I .

Les individus non tirés constituent l'échantillon de base.

On construit une règle de classement à partir de l'échantillon de base. On applique cette règle aux individus de l'échantillon-test : on classe ces individus. Or on connaît leur classement réel. On calcule alors le *pourcentage d'individus bien classés* : ceci est une mesure de la qualité de la règle.

Remarque. On peut utiliser le même principe pour déterminer la valeur optimale du réel r dans la méthode de classement par voisinage (r est le nombre de voisins) ou par boule (r est le rayon de la boule) : on fait varier la valeur de r et on calcule pour chaque valeur le pourcentage de bien classés.

b) Validation croisée (cross-validation)

On divise l'échantillon d'apprentissage I en m parties égales. On détermine la règle de classement à partir de la réunion de $(m - 1)$ parties, qui constitue l'échantillon de base. L'échantillon-test est constitué de la $m^{\text{ième}}$ partie. On calcule le pourcentage de bien classés (% BC) dans l'échantillon-test. On répète ceci m fois, en changeant à chaque fois d'échantillon-test. On calcule le % BC moyen.

Cas particulier : le leave-one-out

Soit $n = \text{card } I$. On prend $m = n$.

On enlève un individu de I . On construit la règle de classement à partir des $(n - 1)$ individus restants. On l'applique à l'individu ôté. On répète cette analyse n fois. On calcule le % BC.

c) Le bootstrap

i) On prélève par un tirage au hasard avec remise n ($n = \text{card } I$) individus de I , qui constituent un échantillon bootstrap. À partir de cet échantillon, on définit une règle de classement qu'on applique aux individus de I . On calcule le pourcentage de bien classés.

On effectue ceci B fois. On mesure la qualité de la règle de classement par le pourcentage de bien classés moyen.

Comme la règle de classement est appliquée à des individus qui ont participé à sa définition, on obtient en général une surévaluation du pourcentage de bien classés.

ii) On utilise alors la méthode suivante.

Pour tout individu i de I , on ne considère que les n_i échantillons bootstrap qui ne contiennent pas i . On applique la règle de classement définie sur chacun de ces échantillons à l'individu i . On note alors le nombre nc_i de fois où i est bien classé et on calcule la fréquence $\frac{nc_i}{n_i}$.

On mesure la qualité de la règle de classement par $\frac{1}{n} \sum_{i=1}^n \frac{nc_i}{n_i}$.

3.3 Critère du pourcentage de bien classés après une classification

1) A partir de r caractères, on effectue *une classification des individus de I* par :

- classification ascendante hiérarchique ;
- centres mobiles, nuées dynamiques ;
- k-means et extensions ;
- algorithme *EM*, etc ...

Ces méthodes sont étudiées dans le chapitre sur la classification.

2) On détermine dans chacune des classes la modalité dominante de y et le nombre d'individus qui ont la modalité dominante.

Règle de sélection. *On retient un ensemble de r caractères qui rend maximal le pourcentage d'individus ayant la modalité dominante de leur classe.*

3.4 Critère de l'analyse de la variance multidimensionnelle

On fait l'hypothèse que les vecteurs des mesures des p caractères quantitatifs sur les n individus de I constituent un échantillon i.i.d. d'un vecteur aléatoire \underline{X} défini sur un espace probabilisé (Ω, \mathcal{A}, P) . Ω est divisé en q classes $\Omega_1, \dots, \Omega_q$. On fait en outre l'hypothèse que, pour tout k , la loi conditionnelle de \underline{X} dans Ω_k est $\mathcal{N}(\underline{m}_k, \underline{\Sigma})$, la matrice de covariance $\underline{\Sigma}$ ne dépendant pas de k .

On note \underline{X}_r la restriction de \underline{X} à r dimensions : $\mathcal{L}(\underline{X}_r | \Omega_k) = \mathcal{N}(m_{kr}, \underline{\Sigma}_r)$.

On considère le test d'analyse de la variance multidimensionnelle au seuil α :

$$\begin{cases} H_0 : \underline{m}_{1r} = \dots = \underline{m}_{qr} \\ H_1 : \overline{H}_0. \end{cases}$$

Soit W_r , respectivement T_r , la restriction de la matrice d'inertie intra W , respectivement totale T , à r dimensions. Sous H_0 , la statistique de réalisation $\frac{\det W_r}{\det T_r}$ est distribuée suivant la loi de Wilks de paramètres $n - 1, r, q - 1, \Lambda_{n-1, r, q-1}$. On détermine $\lambda_{n-1, r, q-1}(\alpha)$ tel que $P(\Lambda_{n-1, r, q-1} < \lambda_{n-1, r, q-1}(\alpha)) = \alpha$.

Règle de décision : si $\frac{\det W_r}{\det T_r} < \lambda_{n-1, r, q-1}(\alpha)$, on rejette H_0 ; sinon, on ne rejette pas H_0 .

On en déduit la règle de sélection suivante :

Règle de sélection. *On retient un ensemble de r caractères qui rend $\frac{\det W_r}{\det T_r}$ minimal, à condition qu'il mène au rejet de l'hypothèse H_0 .*

4 Sélection pas à pas de caractères discriminants

La sélection d'un sous-ensemble de r caractères discriminants parmi un ensemble de p caractères, que l'on a étudiée dans le paragraphe précédent, nécessite de faire C_p^r analyses. Le nombre de calculs peut-être trop important. Les méthodes pas à pas permettent de réduire ce nombre. Cependant elles ne donnent pas de façon générale un sous-ensemble optimal. Le principe est de considérer au pas l l'ensemble des caractères discriminants déterminés au pas $(l-1)$ et, selon la méthode, ascendante ou descendante, de lui adjoindre ou de lui ôter un caractère, en utilisant pour la sélection une mesure du pouvoir discriminant d'un ensemble de caractères.

4.1 Méthode ascendante

4.1.1 Algorithme

On décrit le pas 1 et le pas général l .

1^{er} pas : on détermine le caractère $x^{(1)}$ parmi x^1, \dots, x^p qui a le plus grand pouvoir discriminant vis-à-vis de y .

$i^{\text{ème}}$ pas : on adjoint aux caractères $x^{(1)}, \dots, x^{(l-1)}$ déjà introduits le caractère $x^{(l)}$ tel que l -uplet $(x^{(1)}, \dots, x^{(l-1)}, x^{(l)})$ ait le plus grand pouvoir discriminant parmi les l -uplets $(x^{(1)}, \dots, x^{(l-1)}, x^j)$ pour $j \neq (1), \dots, (l-1)$.

4.1.2 Règles d'arrêt

1) Pourcentage de bien classés

A chaque pas, on applique aux individus de I une règle de classement avec les caractères introduits à ce pas et on calcule le pourcentage de bien classés. On peut arrêter l'introduction des caractères lorsqu'il n'augmente plus de façon importante (il peut diminuer) ou lorsqu'il atteint 100 %.

2) Règle d'arrêt probabiliste

On se place sous les hypothèses du modèle probabiliste normal défini dans le paragraphe précédent : $\mathcal{L}(X | \Omega_k) = \mathcal{N}(\underline{m}_k, \Sigma), k = 1, \dots, q$.

Au pas l , on introduit la variable $X^{(l)}$ telle que $\frac{\det W_l}{\det T_l}$ soit minimal. On teste alors si l'introduction de la variable $X^{(l)}$ améliore la discrimination effectuée par les variables $X^{(1)}, \dots, X^{(l-1)}$:

$$\begin{cases} H_0 : E[X^{(l)} | X^{(1)}, \dots, X^{(l-1)}, I_{\Omega_1}] = \dots = E[X^{(l)} | X^{(1)}, \dots, X^{(l-1)}, I_{\Omega_q}] \\ H_1 : \overline{H_0}. \end{cases}$$

I_{Ω_k} désigne l'indicatrice du sous-ensemble Ω_k .

Soit les matrices d'inertie totale T_l (respectivement T_{l-1}) et intra W_l (resp. W_{l-1}) calculées en utilisant les caractères introduits au pas l (respectivement $l-1$).

On considère la statistique de réalisation

$$\theta_l = \frac{\det W_l}{\det T_l} \cdot \frac{\det T_{l-1}}{\det W_{l-1}}.$$

Sous H_0 , $\frac{1 - \theta_l}{\theta_l} \frac{n - (l - 1) - q}{q - 1}$ est distribué suivant la loi de Snedecor $\mathcal{F}(q - 1, n - (l - 1) - q)$

Règle de décision : si $\frac{1 - \theta_l}{\theta_l} \frac{n - (l - 1) - q}{q - 1} > f_{q-l, n-(l-1)-q}(\alpha)$, on rejette H_0 ; sinon on ne rejette pas H_0 .

On arrête l'introduction d'un caractère lorsqu'on ne rejette pas H_0 .

Exemple. Supposons que, à la première étape, ce soit le caractère "taille" qui soit le plus discriminant, puis que lui soit adjoint le caractère "poids" à la deuxième étape.

Si, à taille égale, les poids moyens dans les classes sont les mêmes (hypothèse H_0 vraie) le caractère "poids" n'améliore pas la discrimination des classes effectuée par le caractère "taille". Il n'est donc pas nécessaire de l'introduire. Néanmoins, considéré seul, le caractère "poids" pourrait être discriminant.

4.2 Méthode progressive

On reprend l'algorithme de la méthode ascendante.

À chaque pas, on teste la capacité de discrimination de chacun des caractères déjà introduits en présence des autres. On peut donc éliminer des caractères introduits aux pas précédents.

4.3 Méthode descendante

On part de l'ensemble des caractères (x^1, \dots, x^p) .

À chaque pas, on ôte le caractère tel que l'ensemble des caractères restants ait le plus grand pouvoir discriminant.

5 Discrimination logistique binaire

On fait l'hypothèse que les vecteurs des mesures des p caractères quantitatifs sur les n individus de I constituent un échantillon i.i.d. d'un vecteur aléatoire \underline{X} dans \mathbb{R}^p défini sur un espace probabilisé (Ω, \mathcal{A}, P) .

5.1 Le modèle logistique binaire

Ω est partitionné en deux classes Ω_1 et Ω_2 (on peut étendre cette étude au cas de q classes: il s'agit alors du modèle logistique polytomique).

On fait l'hypothèse que, pour \underline{x} réalisation possible de \underline{X} , on a :

$$P(\Omega_1 | \underline{X} = \underline{x}) = \frac{e^{\beta_0 + \underline{\beta}' \underline{x}}}{1 + e^{\beta_0 + \underline{\beta}' \underline{x}}} \iff P(\Omega_2 | \underline{X} = \underline{x}) = \frac{1}{1 + e^{\beta_0 + \underline{\beta}' \underline{x}}} ,$$

le vecteur $\underline{\beta}$ de \mathbb{R}^{p^*} , dual de \mathbb{R}^p , et le réel β_0 étant des paramètres inconnus.

Si l'on note :

$$\underline{\theta} = \begin{pmatrix} \beta_0 \\ \underline{\beta} \end{pmatrix} \quad \text{et} \quad \underline{z} = \begin{pmatrix} 1 \\ \underline{x} \end{pmatrix},$$

$$\text{on a : } P(\Omega_1 | \underline{X} = \underline{x}) = \frac{e^{\underline{\theta}' \underline{z}}}{1 + e^{\underline{\theta}' \underline{z}}} \iff \ln \frac{P(\Omega_1 | \underline{X} = \underline{x})}{P(\Omega_2 | \underline{X} = \underline{x})} = \underline{\theta}' \underline{z}.$$

La fonction $S(\underline{x}) = \beta_0 + \underline{\beta}' \underline{x} = \underline{\theta}' \underline{z}$ est appelée la fonction - score.

5.2 Estimation des paramètres au sens du maximum de vraisemblance

On note x_i^1, \dots, x_i^p les valeurs respectives des caractères x^1, \dots, x^p pour l'individu i de I , \underline{x}_i le vecteur $(x_i^1, \dots, x_i^p)'$ et y_i la valeur pour l'individu i de l'indicatrice Y de la classe Ω_1 . On dispose de l'échantillon $((\underline{x}_i, y_i), i = 1, \dots, n)$. On a

$$P(Y = y_i | \underline{X} = \underline{x}_i) = \frac{e^{\underline{\theta}' \underline{z}_i y_i}}{1 + e^{\underline{\theta}' \underline{z}_i}}, \quad i = 1, \dots, n.$$

On définit la vraisemblance à maximiser :

$$L(\underline{\theta}) = \prod_{i=1}^n \frac{e^{\underline{\theta}' \underline{z}_i y_i}}{1 + e^{\underline{\theta}' \underline{z}_i}} = \frac{e^{\underline{\theta}' \sum_{i=1}^n \underline{z}_i y_i}}{\prod_{i=1}^n (1 + e^{\underline{\theta}' \underline{z}_i})}$$

$$\ln L(\underline{\theta}) = \underline{\theta}' \sum_{i=1}^n \underline{z}_i y_i - \sum_{i=1}^n \ln (1 + e^{\underline{\theta}' \underline{z}_i}).$$

On calcule les dérivées partielles de $L(\underline{\theta})$ par rapport aux composantes θ^j de $\underline{\theta}$:

$$\frac{\partial}{\partial \theta^j} \ln L(\underline{\theta}) = \sum_{i=1}^n z_i^j \left(y_i - \frac{e^{\underline{\theta}' \underline{z}_i}}{1 + e^{\underline{\theta}' \underline{z}_i}} \right), \quad j = 1, \dots, p+1.$$

On utilise une méthode numérique itérative, comme celle de Newton-Raphson, pour résoudre le système de $(p+1)$ équations à $(p+1)$ inconnues

$$\frac{\partial}{\partial \theta^j} \ln L(\underline{\theta}) = 0, \quad j = 1, \dots, p+1.$$

5.3 Test de la nullité d'un ensemble de coefficients

Soit $\underline{\theta} = \begin{pmatrix} \underline{\theta}_1 \\ \underline{\theta}_2 \end{pmatrix}$, $\dim \underline{\theta}_2 = r$, $\dim \underline{\theta}_1 = p + 1 - r$.

On note dans la suite $L(\underline{\theta}_1, \underline{\theta}_2)$ la vraisemblance $L(\underline{\theta})$.

On considère le test au seuil α :

$$\begin{cases} H_0 : \underline{\theta}_2 = \underline{0} \\ H_1 : \underline{\theta}_2 \neq \underline{0}. \end{cases}$$

Pour $\dim \underline{\theta}_2 = 1$, si H_0 est vraie, ceci signifie que le caractère de coefficient θ_2 dans le modèle logistique n'est pas discriminant. Pour $\dim \underline{\theta}_2 = r$, si H_0 est vraie, ceci signifie que les r caractères correspondants ne sont pas discriminants dans leur ensemble.

On donne ci-dessous le *test du rapport de vraisemblance*. On utilise également d'autres tests : test de Wald, test du score.

Soit $\hat{\underline{\theta}} = \begin{pmatrix} \hat{\underline{\theta}}_1 \\ \hat{\underline{\theta}}_2 \end{pmatrix}$ l'estimation de $\underline{\theta}$ au sens du maximum de vraisemblance. Soit $\hat{\underline{\theta}}_{10}$ l'estimation de $\underline{\theta}_1$ au sens du maximum de vraisemblance lorsque $\underline{\theta}_2 = \underline{0}$. On considère le rapport de vraisemblance :

$$\lambda(\underline{0}) = \frac{\sup_{\underline{\theta}_1} L(\underline{\theta}_1, \underline{0})}{\sup_{\underline{\theta}_1, \underline{\theta}_2} L(\underline{\theta}_1, \underline{\theta}_2)} = \frac{L(\hat{\underline{\theta}}_{10}, \underline{0})}{L(\hat{\underline{\theta}}_1, \hat{\underline{\theta}}_2)}.$$

Sous H_0 , $-2 \ln \lambda(\underline{0}) = 2(\ln L(\hat{\underline{\theta}}_1, \hat{\underline{\theta}}_2) - \ln L(\hat{\underline{\theta}}_{10}, \underline{0}))$ est asymptotiquement distribué suivant la loi χ_r^2 . On en déduit la règle de décision pour n grand:

Règle de décision : si $-2 \ln \lambda(\underline{0}) > a_r(\alpha)$, on rejette H_0 ; sinon, on ne rejette pas H_0 .

Si l'on ne rejette pas H_0 , les caractères correspondant au vecteur de coefficients $\underline{\theta}_2$ ne sont pas significativement discriminants dans leur ensemble.

Si l'on rejette H_0 , $-2 \ln \lambda(\underline{0})$ est une mesure de la capacité de discrimination en présence des autres caractères de l'ensemble des caractères correspondant à $\underline{\theta}_2$.

5.4 Discrimination logistique pas à pas

5.4.1 Algorithme ascendant

On part du modèle sans caractères explicatifs ($\underline{\beta} = \underline{0}$). On estime le paramètre β_0 par $\hat{\beta}_0$ et on calcule la vraisemblance $L(\hat{\beta}_0) = L_0$.

1^{er} pas : pour $j = 1, \dots, p$, on considère le modèle avec le caractère explicatif x^j , dont on calcule la vraisemblance $L(\hat{\beta}_0, \hat{\beta}^j) = L_1^j$. On retient un caractère qui rend maximale L_1^j à condition que $2(\ln L_1 - \ln L_0) > a_1(\alpha)$ avec $L_1 = \max_j L_1^j$ (on rejette alors l'hypothèse de nullité du coefficient du caractère introduit). Soit $x^{(1)}$ ce caractère.

De façon générale :

r^{ième} pas : pour $j \neq (1), \dots, (r-1) \in \{1, \dots, p\}$, on considère le modèle avec les caractères explicatifs $x^{(1)}, \dots, x^{(r-1)}, x^j$, dont on calcule la vraisemblance $L(\hat{\beta}_0, \hat{\beta}^{(1)}, \dots, \hat{\beta}^{(r-1)}, \hat{\beta}^j) =$

L_r^j . On retient un caractère qui rend maximale L_r^j , à condition que $2(\ln L_r - \ln L_{r-1}) > a_1(\alpha)$, avec $L_r = \max_j L_r^j$. Soit $x^{(r)}$ ce caractère.

On arrête lorsqu'on ne peut plus introduire un nouveau caractère.

5.4.2 Algorithme progressif

On utilise l'algorithme précédent, mais à chaque pas, on teste par la méthode précédente la nullité du coefficient de chacun des caractères déjà introduits et on élimine les caractères pour lesquels on ne rejette pas l'hypothèse de nullité.

5.4.3 Algorithme descendant

On part du modèle avec tous les caractères explicatifs. On en estime les paramètres et on calcule la vraisemblance $L(\widehat{\beta}_0, \widehat{\beta}) = L_p$.

1^{er} pas : pour $j = 1, \dots, p$, on considère le modèle sans le caractère explicatif x^j , dont on calcule la vraisemblance L_{p-1}^j . On ôte un caractère qui rend maximale L_{p-1}^j , à condition que $2(\ln L_p - \ln L_{p-1}) \leq a_1(\alpha)$, avec $L_{p-1} = \max_j L_{p-1}^j$ (on ne rejette pas l'hypothèse de nullité du coefficient du caractère ôté). Soit $x^{(p)}$ ce caractère.

De façon générale :

r^{ième} pas : pour $j \neq (p), (p-1), \dots, (p-r+2) \in \{1, \dots, p\}$, on considère le modèle sans les caractères explicatifs $x^{(p)}, \dots, x^{(p-r+2)}, x^j$, dont on calcule la vraisemblance L_{p-r}^j . On ôte un caractère qui rend maximale L_{p-r}^j , à condition que $2(\ln L_{p-r+1} - \ln L_{p-r}) \leq a_1(\alpha)$ avec $L_{p-r} = \max_j L_{p-r}^j$. Soit $x^{(p-r+1)}$ ce caractère.

On arrête lorsqu'on ne peut plus ôter un caractère.

Eléments d'étude des Séparateurs à Vaste Marge (S.V.M.)

Jean-Marie Monnez, Sophie Wantz-Mézières

On mesure sur un individu p caractères (ou variables) quantitatifs x^1, x^2, \dots, x^p ; on souhaite affecter cet individu à une classe parmi deux classes définies à priori. Pour cela, on définit le caractère

$$(0.1) \quad y = \begin{cases} -1 & \text{si l'individu appartient à la première classe} \\ +1 & \text{si l'individu appartient à la seconde classe} \end{cases}$$

Exemple : Considérons x^1, x^2, \dots, x^p les hauteurs de pics dans des spectres obtenus par spectrométrie de masse dans une analyse protéomique. On souhaite prédire si un individu a ou non un cancer, c'est-à-dire l'affecter à l'une des deux classes définies comme celle des individus cancéreux et celle des individus sains.

Pour déterminer une règle de classement, on dispose d'un ensemble d'apprentissage

$I = \{1, 2, \dots, n\}$. Pour l'individu $i \in I$, on dispose des mesures x_i^1, \dots, x_i^p et de la valeur y_i .

On note I_1 l'ensemble des i tels que $y_i = -1$,
et I_2 l'ensemble des i tels que $y_i = +1$.

On note $x_i = (x_i^1, \dots, x_i^p)$. On dispose donc d'un ensemble de couples $((x_i, y_i))_{1 \leq i \leq n}$.

Dans la méthode étudiée, on représente l'individu i par le point $x_i \in \mathbb{R}^p$ et on cherche un hyperplan de séparation entre les $x_i, i \in I_1$ et les $x_i, i \in I_2$, d'équation

$$(0.2) \quad \beta \cdot x + \beta_0 = 0, \quad (\beta \in \mathbb{R}^{n*}).$$

Dans le cas où l'hyperplan sépare les deux classes I_1 et I_2 , on suppose que

$$(0.3) \quad \begin{aligned} & \text{pour } y_i = 1, \quad \beta \cdot x_i + \beta_0 > 0, \\ & \text{et pour } y_i = -1, \quad \beta \cdot x_i + \beta_0 < 0. \end{aligned}$$

On définit la règle de classement pour un nouvel individu sur lequel on a mesuré x :

$$\begin{cases} \text{si } \beta \cdot x + \beta_0 < 0, & \text{on affecte l'individu à la première classe } I_1, \\ \text{sinon,} & \text{on affecte l'individu à la seconde classe } I_2. \end{cases}$$

On doit distinguer deux cas d'étude, le cas séparable où l'hyperplan sépare parfaitement les classes I_1 et I_2 (on suppose dans ce cas que pour

tout $i \in I$, $y_i (\beta \cdot x_i + \beta_0) > 0$; et le cas contraire (pour lequel il existe des i tels que $y_i (\beta \cdot x_i + \beta_0) \leq 0$).

1 Cas séparable

La distance d'un point x_i à l'hyperplan $H : \beta \cdot x + \beta_0 = 0$ est donnée par :

$$(1.4) \quad \frac{|\beta \cdot x_i + \beta_0|}{\|\beta\|}.$$

On utilise le critère suivant de détermination de l'hyperplan :

Critère : On détermine l'hyperplan H qui maximise la distance au point x_i le plus proche de H .

Soit d_m la distance minimale à l'hyperplan H :

$$(1.5) \quad \frac{|\beta \cdot x_i + \beta_0|}{\|\beta\|} \geq d_m, \quad i = 1, \dots, n$$

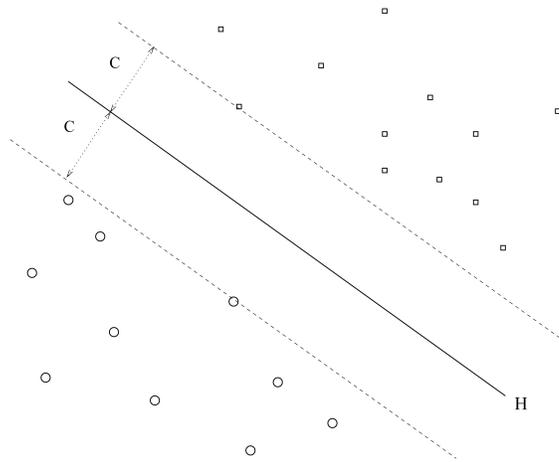


Figure 1.1: Cas séparable ($C = \frac{1}{\|\beta\|} = d_m$)

Donc le problème posé est :

On cherche (β, β_0) qui forment une solution du problème :

$$(1.6) \quad \begin{cases} \text{Max } d_m \\ \text{sous les contraintes } \frac{y_i (\beta \cdot x_i + \beta_0)}{\|\beta\|} \geq d_m, \quad i = 1, \dots, n \end{cases}$$

(on remarque que $y_i(\beta \cdot x_i + \beta_0) = |\beta \cdot x_i + \beta_0|$).

On constate que β et β_0 sont indéterminés à une constante multiplicative près. On va alors poser arbitrairement : $\|\beta\| = \frac{1}{d_m} \Leftrightarrow d_m = \frac{1}{\|\beta\|}$.

On cherche donc (β, β_0) qui forment une solution de :

$$(1.7) \quad \begin{cases} \text{Min } \|\beta\| \\ \text{sous les contraintes } y_i(\beta \cdot x_i + \beta_0) \geq 1, \quad i = 1, \dots, n \end{cases}$$

$$(1.8) \quad \Leftrightarrow \begin{cases} \text{Min } \frac{1}{2} \|\beta\|^2 \\ \text{sous les contraintes } y_i(\beta \cdot x_i + \beta_0) \geq 1, \quad i = 1, \dots, n \end{cases}$$

C'est donc un problème d'optimisation convexe.

On introduit les multiplicateurs de Lagrange $\alpha_i \geq 0$. Soit L la fonction de Lagrange définie par :

$$(1.9) \quad L(\beta, \beta_0; \alpha) = \frac{1}{2} \|\beta\|^2 - \sum_{i=1}^n \alpha_i [y_i(\beta \cdot x_i + \beta_0) - 1]$$

où $\alpha = (\alpha_1, \dots, \alpha_n)$.

On résout le système :

$$(1.10) \quad \begin{cases} \frac{\partial L}{\partial \beta} = \beta - \sum_{i=1}^n \alpha_i y_i x_i = 0 \\ \frac{\partial L}{\partial \beta_0} = - \sum_{i=1}^n \alpha_i y_i = 0 \end{cases}$$

Alors en développant l'expression de L , on obtient :

$$(1.11) \quad \begin{aligned} L(\beta, \beta_0, \alpha) &= \frac{1}{2} \|\beta\|^2 - \sum_{i=1}^n \alpha_i y_i \beta \cdot x_i + \sum_{i=1}^n \alpha_i \\ &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i \cdot x_j \end{aligned}$$

Le problème dual est donc :

$$(1.12) \quad \left\{ \begin{array}{l} \text{Max } \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i \cdot x_j \\ \text{sous les contraintes } \alpha_i \geq 0, \quad i = 1, \dots, n \\ \sum_{i=1}^n \alpha_i y_i = 0 \\ \beta = \sum_{i=1}^n \alpha_i y_i x_i \\ \alpha_i [y_i (\beta \cdot x_i + \beta_0) - 1] = 0, \quad i = 1, \dots, n \end{array} \right.$$

Définition 1.1 On appelle vecteur support tout vecteur x_i tel que $y_i (\beta \cdot x_i + \beta_0) = 1$.

Remarquons qu'un vecteur support est donc un vecteur à distance minimale de l'hyperplan, c'est-à-dire situé sur un des deux bords de la marge. Pour un vecteur support, on a $\alpha_i \geq 0$. Pour tout autre vecteur : $\alpha_i = 0$. Soit S l'ensemble des indices des vecteurs supports ($S \subset I$).

On est donc ramené à un problème de programmation quadratique. Soit une solution de ce problème $(\alpha_1^0, \dots, \alpha_n^0)$. On a alors :

$$(1.13) \quad \begin{cases} \beta = \sum_{i \in S} \alpha_i^0 y_i x_i \\ \beta_0 = \frac{1}{y_i} - \beta \cdot x_i \text{ pour } i \in S. \end{cases}$$

2 Cas non séparable

On se place à présent dans le cas où il n'existe pas d'hyperplan séparant parfaitement les vecteurs x_i de l'ensemble d'apprentissage. Certains points seront mal classés, c'est-à-dire placés du mauvais côté de l'hyperplan.

Pour ces points x_i mal classés, on aura :

$$(2.14) \quad y_i (\beta \cdot x_i + \beta_0) < 0.$$

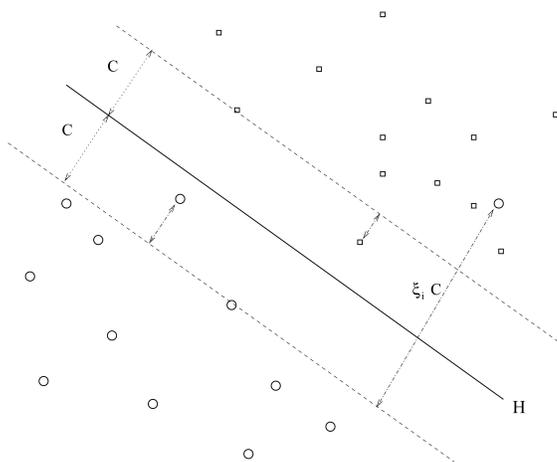


Figure 2.2: Cas non séparable ($C = \frac{1}{\|\beta\|}$)

Le critère dans le cas séparable s'écrit sous la forme :

$$(2.15) \quad \begin{cases} \text{Max} \frac{1}{\|\beta\|} \\ \text{sous les contraintes } y_i(\beta \cdot x_i + \beta_0) \geq 1, \quad i = 1, \dots, n \end{cases}$$

Ceci est maintenant irréalisable. On va alors affaiblir les contraintes en définissant :

$$(2.16) \quad y_i(\beta \cdot x_i + \beta_0) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad i = 1, \dots, n$$

Pour un point x_i mal classé, la distance de ce point à l'hyperplan vérifie :

$$(2.17) \quad \frac{|\beta \cdot x_i + \beta_0|}{\|\beta\|} = \frac{-y_i(\beta \cdot x_i + \beta_0)}{\|\beta\|} \leq (\xi_i - 1) \frac{1}{\|\beta\|}, \quad \text{avec } \xi_i > 1.$$

Pour limiter le nombre de points mal classés et les distances de ces points à l'hyperplan, on cherche à limiter $\sum_{i=1}^n \xi_i$. On définit alors le critère :

$$(2.18) \quad \begin{cases} \text{Min} \left(\frac{1}{2} \|\beta\|^2 + \gamma \sum_{i=1}^n \xi_i \right) \\ \text{sous les contraintes} \quad y_i(\beta \cdot x_i + \beta_0) \geq 1 - \xi_i, \quad i = 1, \dots, n \\ \quad \quad \quad \xi_i \geq 0, \quad i = 1, \dots, n \end{cases}$$

On introduit de nouveaux multiplicateurs $\mu_i \geq 0, i = 1, \dots, n$ et on définit la fonction de Lagrange par :

$$(2.19) \quad L(\beta, \beta_0, \xi; \alpha, \mu) = \frac{1}{2} \|\beta\|^2 + \gamma \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i [y_i(\beta \cdot x_i + \beta_0) - (1 - \xi_i)] - \sum_{i=1}^n \mu_i \xi_i$$

$$(2.20) \quad \text{où } \xi = (\xi_1, \xi_2, \dots, \xi_n), \quad \mu = (\mu_1, \dots, \mu_n).$$

On résout le système :

$$(2.21) \quad \begin{cases} \frac{\partial L}{\partial \beta} = \beta - \sum_{i=1}^n \alpha_i y_i x_i = 0 \\ \frac{\partial L}{\partial \beta_0} = - \sum_{i=1}^n \alpha_i y_i = 0 \\ \frac{\partial L}{\partial \xi_i} = \gamma - \alpha_i - \mu_i = 0, \quad i = 1, \dots, n \end{cases}$$

En développant l'expression de L , on obtient la même expression que dans le cas séparable :

$$(2.22) \quad L(\beta, \beta_0, \xi; \alpha, \mu) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i \cdot x_j$$

Le problème dual est donc :

$$(2.23) \quad \left\{ \begin{array}{l} \text{Max} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i \cdot x_j \\ \text{sous les contraintes} \quad \gamma \geq \alpha_i \geq 0, \sum_{i=1}^n \alpha_i y_i = 0, \mu_i \geq 0, i = 1, \dots, n \\ \xi_i \geq 0, i = 1, \dots, n \\ \mu_i \xi_i = 0, i = 1, \dots, n \\ \alpha_i [y_i (\beta \cdot x_i + \beta_0) - (1 - \xi_i)] = 0, i = 1, \dots, n \\ y_i (x_i \cdot \beta + \beta_0) - (1 - \xi_i) \geq 0, i = 1, \dots, n \end{array} \right.$$

On appelle vecteur support un vecteur x_i tel que $y_i (\beta \cdot x_i + \beta_0) - (1 - \xi_i) = 0$. On a alors $\alpha_i \geq 0$. Pour tout autre vecteur : $\alpha_i = 0$. Il existe deux types de vecteurs supports : ceux pour lesquels on a $\xi_i = 0$, dans ce cas c'est un point bien classé à distance $\frac{1}{\|\beta\|}$ de l'hyperplan, c'est-à-dire sur le bord de la marge, et ceux tels que $\xi_i > 0$, ce sont les points mal classés vérifiant donc $\mu_i = 0$ et $\alpha_i = \gamma$.

Soit S l'ensemble des indices des vecteurs supports, et soit S_0 l'ensemble des indices des vecteurs supports pour lesquels $\xi_i = 0$.

On résout le problème de programmation quadratique convexe (car $\gamma \geq \alpha_i \geq 0$) précédent. Soit une solution $(\alpha_1^0, \dots, \alpha_n^0, \mu_1^0, \dots, \mu_n^0)$. On a alors :

$$(2.24) \quad \left\{ \begin{array}{l} \beta = \sum_{i \in S} \alpha_i^0 y_i x_i \\ \beta_0 = \frac{1}{y_i} - \beta \cdot x_i \text{ pour } i \in S_0. \end{array} \right.$$

En pratique, pour obtenir β_0 , on prend une moyenne de toutes les valeurs obtenues. En effet, les valeurs numériques de β_0 ne sont pas nécessairement strictement égales.

3 Fonction noyau

Au lieu de chercher un hyperplan de séparation optimal dans \mathbb{R}^n , on définit un espace euclidien de représentation intermédiaire (feature space) \mathcal{H} de grande dimension, et une application ϕ de \mathbb{R}^n dans \mathcal{H} .

On cherche alors dans \mathcal{H} un hyperplan de séparation d'équation :

$$(3.25) \quad \beta \cdot \phi(x) + \beta_0 = 0.$$

La solution s'écrit :

$$(3.26) \quad \left\{ \begin{array}{l} \beta = \sum_{i \in S} \alpha_i^0 y_i \phi(x_i) \\ \beta_0 = \frac{1}{y_i} - \beta \cdot \phi(x_i) = \frac{1}{y_i} - \sum_{j \in S} \alpha_j^0 y_j \phi(x_j) \cdot \phi(x_i) \\ \text{pour } i \in S \text{ dans le cas séparable} \\ \text{pour } i \in S_0 \text{ dans le cas non séparable} \end{array} \right.$$

L'équation de l'hyperplan de séparation H s'écrit :

$$(3.27) \quad \sum_{i \in S} \alpha_i^0 y_i \phi(x_i) \cdot \phi(x) + \beta_0 = 0.$$

Les α_i^0 forment une solution du problème :

$$(3.28) \quad \text{Max} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \phi(x_i) \cdot \phi(x_j)$$

sous les contraintes vues précédemment transposées dans \mathcal{H} .

On constate que le problème et sa solution ne dépendent que des produits scalaires $\phi(x) \cdot \phi(x')$. On pose $K(x, x') = \phi(x) \cdot \phi(x')$. K est appelée la fonction noyau. Elle vérifie les propriétés suivantes (conditions de Mercer) :

- * K est symétrique
- * K est semi-définie positive, i.e. pour toute fonction $g \neq 0$ pour laquelle $\int g^2(u) du < \infty$, la condition

$$(3.29) \quad \int \int K(x, x') g(x) g(x') dx dx' \geq 0$$

est vérifiée.

Réciproquement, si les conditions de Mercer sont vérifiées, il existe un espace \mathcal{H} et une fonction ϕ de \mathbb{R}^n dans \mathcal{H} qui vérifient : $K(x, x') = \phi(x) \cdot \phi(x')$.

Plutôt que de choisir l'espace \mathcal{H} et la fonction ϕ , on va alors choisir la fonction K vérifiant les conditions de Mercer.

Exemples de fonction noyau :

$$(3.30) \quad \begin{array}{l} K(x, x') = (1 + x \cdot x')^d \\ K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{c}\right) \\ K(x, x') = \tanh(c_1 x \cdot x' + c_2) \end{array}$$

4 Une méthode de sélection de gènes pour le diagnostic du cancer basée sur les S.V.M., d'après Guyon - Weston - Barnhill - Vapnik *

* *Gene selection for Cancer Classification using Support Vector Machines*, Machine Learning 46, 389-422, 2002.

Les auteurs mettent au point une méthode de sélection préliminaire des gènes avant classification pour éliminer les répétitions et travailler sur des ensembles de gènes plus compacts. Leur méthode est fondée sur les S.V.M. et un processus d'élimination récursif.

Conservons nos notations. Soit l'ensemble d'apprentissage : $I = \{1, 2, \dots, n\}$. Les mesures suivantes sont données : $((x_i, y_i))_{1 \leq i \leq n}$ avec $x_i = (x_i^1, \dots, x_i^p) \in \mathbb{R}^p$, x_i^j étant l'expression du gène j (quantité d'ARNm produite), et $y_i \in \{-1, +1\}$. L'ensemble d'apprentissage est petit : $n \lll p$.

Le principe de la méthode est de ranger les p gènes selon un critère de poids, celui-ci se calculant grâce aux vecteurs α_i obtenus par les S.V.M. Utilisons un algorithme descendant. On démarre avec les p gènes et on en élimine un à chaque étape pour le ranger dans une liste ordonnée. Il y a donc p étapes.

A chaque étape, on applique l'algorithme des S.V.M. aux k gènes restants, ce qui nous fournit une solution $\alpha^0 = (\alpha_1^0, \dots, \alpha_n^0)$. On peut alors calculer le coefficient de chacun des gènes restants dans l'équation de l'hyperplan de séparation :

$$(4.31) \quad \beta_j = \sum_{i=1}^n \alpha_i^0 y_i x_i^j \text{ pour } j = 1, \dots, k$$

On élimine le gène tel que $|\beta_j|$ soit le plus petit. On peut ainsi l'ajouter en tête de la liste des gènes rangés. On s'arrête lorsque tous les gènes ont été éliminés.

On obtient alors une liste des p gènes rangés.

Sélection de composantes pour les SVM (résumé de l'article de Weston *et al.*).

O. Collignon

Données: On dispose d'un ensemble d'apprentissage, représenté par un ensemble de points, $\{(x_i, y_i)_{1 \leq i \leq l}\}$, $x_i \in \mathbb{R}^n$, $y_i = \pm 1$, échantillon de taille l d'une variable aléatoire (X, Y) à valeurs dans $\mathbb{R}^n \times \{-1, 1\}$ et de loi $P_{(X, Y)}$. Par exemple, l'individu i est représenté par les intensités des n biomarqueurs d'un spectre, qui constituent le vecteur x_i , et par son état ($y_i = 1$: cancéreux, $y_i = -1$: sain).

But: Trouver une méthode de sélection d'un sous-ensemble de composantes de X de cardinal m qui discrimine au mieux les deux classes induites par la variable Y .

1 Le problème de sélection des composantes.

Soient $m < n$ et $f : \mathbb{R}^n \rightarrow \{-1, 1\}$ une application qui à un point x de \mathbb{R}^n associe une classe. On définit l'erreur de généralisation $L(f)$:

$$L(f) := E[V(Y, f(X))],$$

où V est une *fonction de perte*, i.e une application

$$V : \{-1, 1\} \times \{-1, 1\} \rightarrow \mathbb{R}_+$$

telle que

$$V(y, y) = 0, \forall y \in \{-1, 1\}.$$

Une telle fonction sert à pénaliser un mauvais classement de l'individu sur lequel l'observation x est faite.

Exemple: $L(f) = P(Yf(X) \leq 0) = E[\mathbb{1}_{\{Yf(X) \leq 0\}}]$ [1]

Dans ce cas, $V(y, z) = \mathbb{1}_{\{yz \leq 0\}}$.

Un autre choix possible est $V(y, z) = \|y - z\|^2$. \square

Le problème est de trouver m composantes de X qui minimisent l'erreur de généralisation $L(f)$.

Lorsqu'on ne connaît pas la loi $P_{(X,Y)}$, on ne peut pas résoudre le problème précédent; on peut alors minimiser l'erreur de généralisation empirique $\hat{L}(f)$:

$$\hat{L}(f) := \frac{1}{l} \sum_{i=1}^l V(y_i, f(x_i)).$$

Dans cet article, on suppose que f appartient à une famille paramétrée de fonctions $\mathcal{F} = \{f(\cdot, \alpha)\}$. Soit $\sigma = (\sigma_1, \dots, \sigma_n) \in \{0, 1\}^n$ et $x * \sigma = (x_1\sigma_1, \dots, x_n\sigma_n)$. L'objectif est de trouver α et σ tels que

$$E[V(Y, f((X * \sigma), \alpha))] \tag{1}$$

soit minimal.

On minimise alors en α et σ l'application

$$\begin{aligned} \tau(\alpha, \sigma) &= E[V(Y, f((X * \sigma), \alpha))] \\ &= \int_{\mathbb{R}^n \times \{-1, 1\}} V(y, f((x * \sigma), \alpha)) dP_{(X,Y)}(x, y) \end{aligned} \tag{2}$$

où

- V est une fonction de perte,
- $\|\sigma\|_0 = m =$ nombre fixé de composantes non nulles de σ ,
- $P_{(X,Y)}$ est inconnue.

La solution peut être approchée en minimisant

$$\tau_{wrap}(\alpha) = \min_{\sigma \in \{0,1\}^n} \tau_{alg}(\sigma), \tag{3}$$

où $\tau_{alg}(\sigma)$ est la sortie d'un algorithme d'apprentissage d'entrée σ . La méthode présentée dans la suite tire partie des performances de cette procédure, tout en diminuant la complexité des calculs.

2 Rappels sur les SVM.

On suppose que les deux classes de points $(x_i)_{1 \leq i \leq l}$ de \mathbb{R}^n sont séparables par un hyperplan.

On cherche un hyperplan H d'équation:

$$H : f(x) = w \cdot x + b = 0.$$

On suppose que pour $y_i = 1$, $w \cdot x_i + b > 0$ et pour $y_i = -1$, $w \cdot x_i + b < 0$. Dans ce cas, $y_i(w \cdot x_i + b) = |w \cdot x_i + b|$.

La distance d'un point x_i à l'hyperplan H est

$$\frac{|w \cdot x_i + b|}{\|w\|}. \quad (4)$$

On définit la notion de marge de l'hyperplan H :

$$\text{marge}(H) = \min_{1 \leq i \leq l} \frac{y_i(w \cdot x_i + b)}{\|w\|}.$$

L'hyperplan cherché est celui ayant la marge maximale.

Trouver cet hyperplan revient à résoudre le problème de minimisation sous contraintes:

$$\begin{cases} \min \frac{1}{2} \|w\|^2 \\ y_i(w \cdot x_i + b) \geq 1, \forall i = 1, \dots, l. \end{cases} \quad (5)$$

Problème dual. Des multiplicateurs de Lagrange $(\alpha_i)_{1 \leq i \leq l}$ sont introduits pour chaque contrainte. Le problème dual s'écrit:

$$\begin{cases} \max W^2(\alpha) = \max \left(\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{1 \leq i, j \leq l} \alpha_i \alpha_j y_i y_j x_i \cdot x_j \right) \\ \alpha_i \geq 0, \forall i = 1, \dots, l \\ \sum_{i=1}^l \alpha_i y_i = 0 \\ \alpha_i (y_i (w \cdot x_i + b) - 1) = 0, \forall i = 1, \dots, l. \end{cases} \quad (6)$$

On résout ce problème. Soit $(\alpha_1^0, \dots, \alpha_l^0)$ une solution de ce problème. On a

$$w^0 = \sum_{i=1}^l \alpha_i^0 y_i x_i.$$

Les α_i^0 non nuls correspondent aux points les plus proches de l'hyperplan; on les appelle *vecteurs supports*.

De là est déduite l'équation de l'hyperplan de séparation:

$$f(x) = \sum_{i=1}^l \alpha_i^0 y_i x_i \cdot x + b = 0.$$

Pour un individu sur lequel on a mesuré x , on applique la règle de décision suivante:

- si $f(x) > 0$, on affecte l'individu à la classe $\{y = 1\}$,
- si $f(x) < 0$, on affecte l'individu à la classe $\{y = -1\}$.

Fonction noyau. Au lieu de chercher un hyperplan dans \mathbb{R}^n , on définit un espace euclidien de représentation intermédiaire (*feature space*) \mathcal{H} de grande dimension, et une application $\phi : \mathbb{R}^n \rightarrow \mathcal{H}$.

On cherche alors un hyperplan de séparation d'équation:

$$w \cdot \phi(x) + b = 0.$$

La fonction $f(x)$ s'écrit alors

$$f(x) = w^0 \cdot \phi(x) + b = \sum_{i=1}^l \alpha_i^0 y_i \phi(x_i) \cdot \phi(x) + b. \quad (7)$$

$\alpha_1^0, \dots, \alpha_l^0$ forment une solution du problème:

$$\left\{ \begin{array}{l} \max W^2(\alpha) = \max \left(\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{1 \leq i, j \leq l} \alpha_i \alpha_j y_i y_j \phi(x_i) \cdot \phi(x_j) \right) \\ \alpha_i \geq 0, \forall i = 1, \dots, l \\ \sum_{i=1}^l \alpha_i y_i = 0 \\ \alpha_i (y_i (w \cdot \phi(x_i) + b) - 1) = 0, \forall i = 1, \dots, l. \end{array} \right. \quad (8)$$

Le problème et sa solution ne dépendent que des produits scalaires $\phi(x) \cdot \phi(x')$.
On pose alors:

$$K(x, x') = \phi(x) \cdot \phi(x')$$

Plutôt que de définir \mathcal{H} et ϕ , on définit ainsi une fonction $K : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, appelée *fonction noyau*.

La fonction K est symétrique:

$$K(x, x') = K(x', x), \quad \forall x, x' \in \mathbb{R}^n,$$

Conditions de Mercer. Une application K symétrique est un noyau si pour tous les x_i possibles, la matrice M d'élément (i, j)

$$K(x_i, x_j)$$

est semi-définie positive, i.e

$$x' M x \geq 0, \quad \forall x \in \mathbb{R}^n.$$

On montre que si une fonction K vérifie les conditions de Mercer, il existe un espace \mathcal{H} et une application ϕ telle que:

$$K(x, x') = \phi(x) \cdot \phi(x') \quad \square$$

Théorème. Si pour chaque ensemble d'apprentissage de cardinal l , les images $\{\phi(x_i)\}_{1 \leq i \leq l}$ sont séparables par un hyperplan de \mathcal{H} et sont à l'intérieur d'une sphère de rayon R , et si M désigne la marge maximale, alors la valeur moyenne sur les ensembles d'apprentissage de la probabilité d'erreur de classement est majorée par

$$\frac{1}{l} E\left[\frac{R^2}{M^2}\right] = \frac{1}{l} E[R^2 W^2(\alpha^0)]. \quad (9)$$

Remarque: R et M sont des variables aléatoires [2].

3 Sélection de composantes pour les SVM.

Soit une application ϕ fixée. On veut trouver la fonction

$$f(x; w, b, \sigma) = w \cdot \phi(x * \sigma) + b$$

qui minimise l'erreur de généralisation $E[V(Y, f(X; w, b, \sigma))]$.

On pose

$$\phi_\sigma(x) = \phi(x * \sigma)$$

et

$$K_\sigma(x, y) = K((x * \sigma), (y * \sigma)) = \phi_\sigma(x) \cdot \phi_\sigma(y). \quad (10)$$

Le théorème du paragraphe précédent reste vrai avec ces noyaux.

D'après l'équation (9), on minimise

$$\tau_{alg}(\sigma) = R^2(\sigma)W^2(\alpha^0, \sigma), \quad (11)$$

où le rayon pour le noyau K_σ peut-être calculé [2] par

$$R^2(\sigma) = \max_{\beta} \left(\sum_{1 \leq i \leq l} \beta_i K_\sigma(x_i, x_i) - \sum_{1 \leq i, j \leq l} \beta_i \beta_j K_\sigma(x_i, x_j) \right), \quad (12)$$

sous les contraintes

$$\begin{cases} \sum_{i=1}^l \beta_i = 1, \\ \beta_i \geq 0, \forall i = 1, \dots, l \end{cases} \quad (13)$$

et où $W^2(\alpha^0, \sigma)$ est défini par le maximum de

$$\sum_{1 \leq i \leq l} \alpha_i - \frac{1}{2} \sum_{1 \leq i, j \leq l} \alpha_i \alpha_j y_i y_j K_\sigma(x_i, x_j). \quad (14)$$

Au lieu de minimiser $R^2(\sigma)W^2(\alpha^0, \sigma)$ pour $\sigma \in \{0, 1\}^n$, ce qui nécessite de considérer tous les sous-ensembles de n éléments égaux à 0 ou 1, on suggère la méthode suivante.

On minimise pour $\sigma \in \mathbb{R}^n$, et on ajoute une pénalisation pour forcer σ à rester proche de $B = \{\sigma \in \{0, 1\}^n, \|\sigma\|_0 = m\}$. On cherche donc:

$$\min_{\sigma \in \mathbb{R}^n} \left([R^2W^2](\sigma) + \lambda \sum_{i=1}^l \sigma_i^p \right),$$

sous les contraintes

$$\begin{cases} \sum_{i=1}^l \sigma_i = m, \\ \sigma_i \geq 0, \forall i = 1, \dots, l. \end{cases} \quad (15)$$

Pour cela, on détermine:

$$\frac{\partial R^2 W^2(\sigma)}{\partial \sigma_k} = R^2(\sigma) \frac{\partial W^2(\alpha^0, \sigma)}{\partial \sigma_k} + W^2(\alpha^0, \sigma) \frac{\partial R^2(\sigma)}{\partial \sigma_k} \quad (16)$$

$$\frac{\partial R^2(\sigma)}{\partial \sigma_k} = \sum_{i=1}^l \beta_i^0 \frac{\partial K_\sigma(x_i, x_i)}{\partial \sigma_k} - \sum_{1 \leq i, j \leq l} \beta_i^0 \beta_j^0 \frac{\partial K_\sigma(x_i, x_j)}{\partial \sigma_k} \quad (17)$$

$$\frac{\partial W^2(\alpha^0, \sigma)}{\partial \sigma_k} = - \sum_{1 \leq i, j \leq l} \alpha_i^0 \alpha_j^0 y_i y_j \frac{\partial K_\sigma(x_i, x_j)}{\partial \sigma_k} \quad (18)$$

et on résout l'équation

$$\frac{\partial R^2 W^2(\sigma)}{\partial \sigma_k} = 0, \forall k = 1, \dots, n$$

avec (16) dans laquelle on reporte (17) et (18).

Bibliographie

- [1] O.Bousquet, *Introduction aux "Support Vector Machine"*, 2001
<http://www.math.u-psud.fr/blanchard/gtsvm/intro.pdf>.
- [2] V.Vapnik, *Statistical Learning Theory*. John Wiley and Sons, New York, 1998.
- [3] J.Weston *et al.*, *Feature Selection for SVMs*, Advances in Neural Information Processing Systems, 13, SOLLA, S.A. , LEEN, T.K. , MULLER, K.-R., Eds. MIT Press, Cambridge, MA, pp.668-674, 2001.

Méthodes statistiques mises en oeuvre dans “Three Biomarkers Identified From Serum Proteomic Analysis for the Detection of Early Stage Ovarian Cancer. ” de Z.Zhang

O. Collignon

Le but de l'article est d'identifier dans le sérum des biomarqueurs du cancer des ovaires. Pour cela, une analyse protéomique est réalisée sur cinq sites de différents pays, afin de comparer les résultats obtenus et de valider de manière indépendante les éventuels biomarqueurs (c'est-à-dire de mesurer la capacité qu'ont ces biomarqueurs à détecter le cancer des ovaires). La détection des biomarqueurs a été réalisée grâce aux échantillons des deux premiers sites; les trois biomarqueurs trouvés ont ensuite été validés avec les échantillons des troisième et quatrième sites principalement. Enfin, des tests statistiques ont été réalisés sur ces biomarqueurs à partir des échantillons du cinquième site, afin d'obtenir une confirmation indépendante. Tout ce protocole a permis d'identifier trois biomarqueurs.

Pour cette étude, 645 échantillons de sérum ont été sélectionnés: 503 sur les quatre premiers sites et 142 pour le cinquième. Ces échantillons ont été prélevés sur des patients atteints de cancers des ovaires à différents stades et de différents types, ainsi que sur des patients sains (les contrôles).

1 Identification des pics informatifs.

Pour minimiser les risques de découvrir de faux biomarqueurs, la recherche des pics informatifs a été réalisée indépendamment sur deux sites différents. Pour cette étude, 20 échantillons cancéreux aux stades I/II, 30 échantillons sains et 50 échantillons cancéreux bénins ont été prélevés sur le premier site, contre 37 échantillons cancéreux aux stades I-IV, 49 échantillons sains et 90 échantillons cancéreux bénins sur le deuxième. Pour la recherche de biomarqueurs potentiels, seuls les échantillons cancéreux aux stades I-IV et sains de

chaque site ont été analysés (constituant les ensembles appelés discovery set 1 et 2).

L'analyse protéomique de cette étude est sensiblement la même que celle réalisée pour la recherche de biomarqueurs du cancer du colon: les protéines des échantillons sont séparées en fonction de leur pH, puis fixées sur des barrettes possédant différentes propriétés. Elles sont ensuite analysées par un spectromètre de masse CIPHERGEN (modèle précédent de celui utilisé pour la recherche de biomarqueurs du cancer du colon), qui fournit des séries de profils.

Remarque: Le sommet d'un pic d'un profil est défini graphiquement comme étant un point vérifiant $\frac{\text{signal}}{\text{bruit}} > 5$ (Fig.1); cette valeur seuil peut-être diminuée afin de faire augmenter le nombre de pics et affiner l'étude.□

A l'issue des deux analyses de chaque site, seuls les biomarqueurs statistiquement informatifs et ayant les mêmes caractéristiques irrégulières (hauteur, largeur à mi-hauteur) dans les deux discovery sets ont été conservés comme biomarqueurs potentiels, formant l'ensemble cross-validation. Ainsi, trois biomarqueurs ont été identifiés.

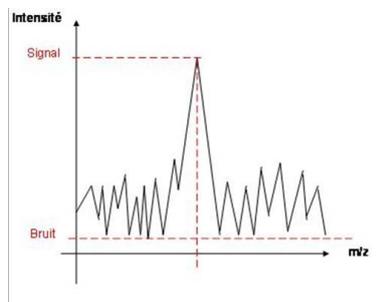


Figure 1: Le rapport signal/bruit.

Méthodes statistiques:

- L'algorithme *Unified Maximum Separability Analysis (UMSA)* a été utilisé pour analyser les intensités des pics et choisir les pics informatifs. Cet algorithme initialise le *Support Vector Machine algorithm (SVM)* avec des données de méthodes statistiques multivariées. Ces méthodes semblent jouer un rôle déterminant dans les analyses de Zhang et doivent donc être étudiées.

- La technique du *Bootstrap* est également appliquée pour identifier les pics ayant un fort pouvoir discriminant sur de multiples sous-populations ré-échantillonnées.

2 Comparaison du pouvoir discriminant de chacun des trois biomarqueurs et de CA125.

L'étape suivante est d'évaluer sur des patients la capacité des biomarqueurs identifiés et de la protéine CA125 à détecter le cancer des ovaires. Le CA125 est une protéine présente en forte quantité dans le sérum de 80 % des femmes atteintes du cancer des ovaires, au moment du diagnostic de la maladie. Il semble donc intéressant de faire intervenir cette protéine dans l'analyse. Pour cette étude, les ensembles discovery set 1 et 2 sont combinés. L'ensemble validation set 1 est de plus constitué à partir des échantillons restants des deux premiers sites (cancéreux bénins), ainsi que des échantillons des troisième et quatrième sites. Tout ces échantillons sont de types cancéreux aux stades I-IV, sains et cancéreux bénins.

2.1 Comparaisons contrôles/cancéreux.

L'idée est de confronter les patients sains aux patients cancéreux en comparant leurs quantités de chaque biomarqueur et de CA125 dans le sérum. Ainsi, si les biomarqueurs identifiés sont fiables, leurs quantités sérumiques devraient être significativement différentes d'un groupe à l'autre. La table 2 de l'article présente les statistiques descriptives (moyenne \pm écart-type, médiane) ainsi que les distributions des quantités de chacun des trois biomarqueurs et de CA125, dans le sérum des individus des groupes discovery sets 1+2 et validation set 1. Dans chaque groupe, la distribution empirique de la quantité d'un biomarqueur chez les patients sains est comparée à cette même distribution chez les divers types de cancéreux. Cette comparaison est réalisée à l'aide du test de *Mann-Whitney*.

2.2 Comparaisons par groupes.

Les distributions des trois biomarqueurs sont cette fois comparées par sous-groupe d'échantillons cancéreux (stade du cancer, type histologique (type de tissu) du cancer, âge du patient), afin de mesurer par exemple l'influence de l'âge sur le développement d'un cancer. Le test de *Mann-Whitney* a été utilisé pour les comparaisons des quantités des biomarqueurs entre deux groupes, et

le test de *Kruskal-Wallis* pour les comparaisons entre plusieurs groupes. Les résultats sont présentés dans un tableau avec les statistiques descriptives.

2.3 Comparaison du pouvoir discriminant de chacun des trois biomarqueurs à celui de CA125:

Il semble maintenant intéressant de voir si les biomarqueurs identifiés détectent le cancer des ovaires avec une plus grande efficacité que le CA125. Le pouvoir discriminant est la capacité qu'a un marqueur à détecter le cancer des ovaires. Il est mesuré avec la spécificité (capacité à exclure un patient sain) et la sensibilité (capacité à détecter un cancer) d'un test cancer/non cancer propre à chaque biomarqueur. Plus le test est spécifique et sensible, plus il est efficace. Chaque test, réalisé ici sur les individus des discovery sets 1 + 2 et validation set 1, dépend d'un seuil s^* : tous les individus se voient attribuer un score s (calculé semble-t-il grâce à un *modèle logistique*); le test d'un individu est dit positif si $s > s^*$ et négatif si $s \leq s^*$. La spécificité* et la sensibilité* du test sont alors calculées avec les scores de tous les individus. En faisant varier ce seuil, les couples (1-spécificité,sensibilité) obtenus permettent de tracer une courbe, appelée *courbe de ROC*. Plus l'aire sous la courbe est proche de 1, plus le test est efficace. Des figures permettent de comparer le pouvoir discriminant de chacun des trois biomarqueurs à celui de CA125 (figures différentes selon le type de cancer et l'ensemble d'individus testé), en présentant les courbes de ROC correspondantes.

(*): Pour les définitions mathématiques de la spécificité et de la sensibilité d'un test, voir le résumé de l'article sur la triangulation moléculaire (Actes 2005).

2.4 Pouvoir discriminant des trois biomarqueurs ensemble et des trois biomarqueurs + CA125:

L'étape suivante est de regarder si la combinaison des trois biomarqueurs (avec et sans CA125) discrimine plus efficacement les cancéreux des non-cancéreux, que le CA125 seul. Pour cela, les deux ensembles utilisés pour l'identification des biomarqueurs sont réunis puis séparés aléatoirement en un ensemble d'apprentissage et un ensemble de test. Les données disponibles pour chaque individu de ces ensembles sont les intensités des pics des trois biomarqueurs et de CA125. Deux *modèles prédictifs multivariés non linéaires* sont construits:

1. Le premier n'utilise que les données relatives aux trois biomarqueurs;

2. Le second utilise les mêmes données plus celles du CA125.

L'idée est de comparer le pouvoir discriminant des trois biomarqueurs ensemble, des trois biomarqueurs et du CA125 ensemble, et du CA125 seul.

Pour cela:

- de nouvelles courbes de ROC sont tracées à partir des données de différents ensembles;
- en utilisant les seuils qui maximisent la somme de la sensibilité et de la spécificité sur l'ensemble d'apprentissage, les pouvoirs discriminants de ces modèles ont été évalués sur le test set et le validation set 1. Pour cela, les sensibilités et les spécificités des deux modèles et du CA125 ont été comparées. D'après l'article, ces comparaisons utilisent de nouveau l'UMSA (non linéaire). Les résultats sont présentés dans un tableau.

3 Validation indépendante des biomarqueurs.

Pour une validation indépendante des biomarqueurs identifiés, les distributions des quantités de CA125 et des trois biomarqueurs dans le sérum des 142 échantillons du cinquième site sont comparées. Ces distributions sont comparées avec le test de Mann-Whitney. Un tableau présente les résultats correspondants.

Bibliographie

- [1] Z.Zhang *Three Biomarkers Identified From Serum Proteomic Analysis For The Detection Of Early Stage Ovarian Cancer*. Cancer Research 64, 5882-5890, 2004.

Méthodes mises en oeuvre dans “Use of proteomic pattern in serum to identify ovarian cancer. ” de E.F. Petricoin et “Reproducibility of SELDI-TOF protein patterns in serum... ” de K.A. Baggerly.

O. Collignon

Le but de l'article de Petricoin est de déterminer dans le sérum des biomarqueurs du cancer des ovaires par une analyse protéomique, via une méthode de classement. L'article de Baggerly s'intéresse dans un premier temps à la reproductibilité des résultats de Petricoin, puis propose une autre méthode pour déterminer des biomarqueurs.

1 La méthode de Petricoin.

Dans cette étude, 216 individus ont été recrutés: 100 sains, 100 malades et 16 individus ayant des symptômes bénins (dataset 1). Grâce à un spectromètre de masse, chaque individu fournit un unique spectre.

Chacun des 216 spectres est composé d'environ 15000 points, et est représenté par un vecteur $S^j = \{(x_i, y_i^j)_{1 \leq i \leq 15000}\}$, $1 \leq j \leq 216$, où $x_i := \frac{m_i}{z_i}$ est un rapport masse/charge de protéines, et y_i^j est l'intensité correspondante pour le j ème spectre (normalisée). Les x_i sont fixes et sont les mêmes pour tous les spectres. La recherche des biomarqueurs est ici effectuée sur tous les points du spectre plutôt que sur les pics uniquement.

Le traitement des données est mené en deux étapes:

1. l'identification des biomarqueurs, en utilisant le critère du pourcentage de bien-classés après une classification, réalisée sur un ensemble d'apprentissage composé de 50 spectres sains et 50 spectres cancéreux,

2. la confirmation des biomarqueurs par une méthode de classement d'un ensemble de test, composé de 50 spectres sains, 50 spectres cancéreux et 16 spectres bénins.

1.1 Identification des biomarqueurs.

Un ensemble d'apprentissage est créé à partir de 50 spectres sains et 50 spectres cancéreux afin d'identifier les biomarqueurs. Le but est de trouver un sous-ensemble $e^* = \{x_{i_1}, \dots, x_{i_n}\}$ de $\{x_1, \dots, x_{15000}\}$ qui discrimine le mieux les spectres cancéreux des spectres non-cancéreux.

Pour cela, 1500 sous-ensembles de $\{x_1, \dots, x_{15000}\}$ de cardinaux compris entre 5 et 20 sont choisis aléatoirement:

$$e_i = \{x_{i,1}, \dots, x_{i,n_i}\}, \quad 1 \leq i \leq 1500, \quad 5 \leq n_i \leq 20.$$

Afin de trouver le sous-ensemble le plus discriminant, un score est calculé pour chaque sous-ensemble e_i , de sorte que ce score augmente avec le pouvoir discriminant des sous-ensembles. Pour cela, une classification de tous les spectres est réalisée à i fixé.

Classification des spectres. Pour chacun des 100 individus, on regarde les ensembles d'intensités correspondant aux masses du sous-ensemble $e_i = \{x_{i,1}, \dots, x_{i,n_i}\}$:

$$f_i^j = \{y_{i,1}^j, \dots, y_{i,n_i}^j\}, \quad 1 \leq j \leq 100.$$

L'idée est ensuite de répartir ces ensembles en différentes classes, en fonction de leurs distances relatives. On procède pas-à-pas.

- On commence par f_i^1 et f_i^2 :

si

$$d(f_i^1, f_i^2) := \sqrt{\sum_{k=1}^{n_i} |y_{i,k}^1 - y_{i,k}^2|^2} \leq \frac{\sqrt{n_i}}{10},$$

alors f_i^1 et f_i^2 sont placés dans la même classe. On calcule ensuite le vecteur moyen $z_i^1 = \frac{1}{2}(f_i^1 + f_i^2)$ de ces deux vecteurs, que l'on appelle *centre* de la première classe;

sinon,

on place f_i^1 et f_i^2 dans deux classes différentes, dont les centres sont définis respectivement par $z_i^1 = f_i^1$ et $z_i^2 = f_i^2$.

- On itère le processus avec f_i^3 :

si f_i^1 et f_i^2 sont placés dans la même classe, on étudie la distance de f_i^3 à z_i^1 comme précédemment ; si $d(f_i^3, z_i^1) \leq \frac{\sqrt{n_i}}{10}$, le centre de la classe sera alors remplacé par le nouveau vecteur moyen $z_i^1 = \frac{1}{3}(f_i^1 + f_i^2 + f_i^3)$;

si f_i^1 et f_i^2 sont placés dans deux classes différentes: si $d(f_i^3, z_i^1) > \frac{\sqrt{n_i}}{10}$ et $d(f_i^3, z_i^2) > \frac{\sqrt{n_i}}{10}$, on place f_i^3 dans une nouvelle classe dont le centre z_i^3 vaut f_i^3 ; sinon on affecte f_i^3 à la classe dont il est le plus proche du centre.

La procédure s'arrête lorsque tous les spectres ont été affectés à une classe. \square

Calculs des scores. Une fois tous les spectres regroupés en différentes classes, un score est calculé pour chaque sous-ensemble e_i .

La nature des spectres (cancer/non-cancer) étant connue, une classe est déclarée “ cancéreuse ” si elle contient une majorité de spectres cancéreux; à l'inverse, une classe est déclarée “ non-cancéreuse ” si elle contient une majorité de spectres sains. Le nombre de bien-classés est alors compté dans chaque classe; le pourcentage total de bien classés s_i est appelé score de l'ensemble e_i . \square

Les ensembles e_i ayant les plus grands scores sont conservés et réunis dans un même nouvel ensemble. L'auteur ne précise pas combien d'ensembles sont conservés. Le processus est alors répété depuis le début sur cet ensemble (sélection aléatoire de sous-ensembles, classement des spectres, calcul des scores). La procédure s'arrête lorsqu'il ne reste plus qu'un seul sous-ensemble discriminant e^* de cardinal 5. D'après l'article de Petricoin, la procédure a permis d'identifier 5 biomarqueurs au bout de 250 itérations.

1.2 Classement des spectres de l'ensemble de test.

Pour chacun des 116 spectres de l'ensemble de test, on regarde les ensembles d'intensités correspondant au sous-ensemble de masses e^* , que l'on note $f_*^j = \{y_{i_1}^j, \dots, y_{i_n}^j\}$, $1 \leq j \leq 116$, ($n = \text{card}(e^*) = 5$).

A la fin de l'étape précédente, le sous-ensemble e^* a induit m classes, toutes annotées “ cancéreuse ” ou “ non-cancéreuse ”. Les centres de ces classes sont notés $\{z_*^1, \dots, z_*^m\}$. Le but est maintenant de classer chacun des 116 spectres de l'ensemble de test dans l'une de ces m classes. Les bien classés seront alors comptés pour évaluer la pertinence des biomarqueurs identifiés.

Ceci permettra d'évaluer le pouvoir discriminant de e^* et de confirmer les biomarqueurs correspondants.

Pour chaque spectre f_*^j , la procédure est la suivante:

si $d(f_*^j, z_*^k) > \frac{\sqrt{n^*}}{10}$, $\forall 1 \leq k \leq m$, f_*^j est placé dans une nouvelle classe appelée "autres", de centre $z_*^{m+1} = f_*^j$;

sinon, f_*^j est affecté à la classe dont il est le plus proche du centre. Le centre est alors remplacé par la moyenne des éléments de la classe.

Lorsque tous les individus sont classés, les classes sont réunies par types "cancéreuses", "non-cancéreuses" et "autres", en utilisant un étiquetage avec la règle de la majorité, comme précédemment. L'article de Petricoin conclut que 50 spectres cancéreux sur 50 et 47 spectres sains sur 50 ont été classés correctement; de plus 16 spectres bénins sur 16 ont été classés comme "autres".

2 Les trois études de Baggerly.

2.1 Reproductibilité de la méthode de Petricoin.

Baggerly propose dans la première de ses trois études de vérifier la reproductibilité des résultats de Petricoin, en réutilisant la méthode que nous venons de décrire sur le même ensemble de données. Il n'a pas obtenu les mêmes résultats, ceci étant dû selon lui au fait que Petricoin aurait traité des données brutes, tandis que Baggerly a testé sa méthode sur des données normalisées.

Baggerly a tout d'abord calculé une matrice de distance M définie comme suit:

$$M = (d(f_*^j, f_*^l))_{1 \leq j, l \leq 216}.$$

Il a alors noté une absence de distances importantes entre les spectres de même type et les spectres de types différents. En effet, si le critère de distance choisi était idéal, les distances entre spectres de même type seraient faibles, tandis que les distances entre spectres de types différents seraient élevées.

Par ailleurs, seules quatre distances de la matrice sont strictement supérieures à la distance fixée arbitrairement à $\frac{\sqrt{5}}{10}$, soit la borne utilisée dans l'algorithme pour créer une nouvelle classe, lorsqu'on travaille avec des ensembles de cardinal 5. Autrement dit, la méthode a créé deux classes, l'une de 212 spectres

et l'autre de 4 spectres. Ces résultats sont clairement moins bons que ceux de Petricoin.

Enfin, Baggerly émet une série de remarques:

- le choix du cardinal de l'ensemble discriminant e^* est arbitrairement fixé à cinq; il semblerait tout aussi pertinent de chercher un sous-ensemble de cardinal plus petit ;
- le pouvoir discriminant de chaque ensemble e_i n'est mesuré qu'à partir de la qualité du classement qu'il induit ;
- il est possible d'utiliser une autre distance que la distance euclidienne, qui tiendrait compte de la dispersion des classes ;
- la création d'un grand nombre de classes à trop peu d'éléments devrait être pénalisée ;
- la borne choisie $\frac{\sqrt{n_i}}{10}$ est arbitraire ;
- comment choisir le meilleur sous-ensemble, s'il existe deux sous-ensembles e_1^* et e_2^* qui classent parfaitement tous les spectres?

2.2 Généralisation des résultats.

Dans sa deuxième étude, Baggerly souhaite voir s'il est possible de classer un ensemble d'individus à l'aide de biomarqueurs identifiés sur un autre ensemble. Il dispose pour cela de deux ensembles de données, l'un correspondant aux mêmes 216 spectres mais traités avec un autre spectromètre de masse (dataset 2), et un autre ensemble contenant 253 spectres (91 normaux et 162 cancéreux), obtenus à partir du même spectromètre de masse que le dataset 2 (dataset 3).

La méthode précédente a permis à Baggerly de trouver 7 autres biomarqueurs sur le dataset 3. Pour généraliser les résultats, les biomarqueurs trouvés pour un ensemble de données devraient également discriminer l'autre ensemble de données. Pour cela, une première matrice de distances est calculée pour le dataset 3 comme précédemment à partir des 7 biomarqueurs correspondants; via un code de couleur la séparation des spectres sains et cancéreux apparaît clairement. En revanche, en utilisant les 5 biomarqueurs du dataset 2 (qui sont les mêmes que ceux du dataset 1) pour calculer une matrice de distance sur le dataset 3, les spectres ne sont pas séparés correctement. Ceci pose le problème de la généralisation des résultats pour cette méthode.

2.3 Un autre moyen de confirmer les biomarqueurs.

Afin de confirmer les 7 biomarqueurs du dataset 3, Baggerly propose d'utiliser un test de Student de comparaison de moyennes. En effet, pour un biomarqueur donné, soient X la variable aléatoire ayant pour réalisation l'intensité du biomarqueur chez un patient sain et Y la variable aléatoire ayant pour réalisation l'intensité du même biomarqueur chez un patient cancéreux. Si le biomarqueur considéré discrimine effectivement les spectres cancéreux des spectres non cancéreux, l'hypothèse

$$H_0 : E(X) = E(Y)$$

devrait être rejetée. Pour tester cette hypothèse, à chaque biomarqueur est associée une valeur t , réalisation de la statistique T du test (qui est supposée suivre une loi de Student, voir remarque ci-dessous). Dans le cas où H_0 est rejetée, plus t est élevée, plus le biomarqueur est jugé discriminant. Ainsi le biomarqueur ayant la valeur t la plus élevée a permis de classer correctement 238 des 253 spectres.

Cette méthode peut donc également permettre de détecter des biomarqueurs, en calculant une valeur de t pour tous les points m/z des spectres, et en conservant ceux qui ont la valeur de t la plus élevée. Baggerly a utilisé ce procédé pour trouver des biomarqueurs du dataset 3 différents de ceux de Petricoin. Les biomarqueurs ainsi trouvés classent bien les spectres, mais certains sont situés dans la zone de bruit (en début de spectre, les protéines représentées ne proviennent pas des échantillons sanguins analysés).

Remarque: Pour utiliser un test de Student, X et Y doivent suivre des lois normales et être d'écart-types égaux; néanmoins ces hypothèses ne semblent pas avoir été testées dans l'article.

Bibliographie

- [1] E.F Petricoin III *et al*, *Use of proteomic patterns in serum to identify ovarian cancer*. Lancet; vol.359, 2002, pp. 572-577.
- [2] K.A Baggerly *et al*, *Reproductibility of SELDI-TOF protein patterns in serum: comparing datasets from different experiments*, Bioinformatics, Vol. 20, n°5, 2004, pp 777-785.
- [3] J.T Tou *et al*, Eds. *Pattern recognition principles*, Reading, MA: Addison Wesley Publishing Company, 1974.

L'analyse en composantes principales

Jean-Marie MONNEZ

Contents

1	Données et buts	2
2	Etude géométrique dans l'espace des individus \mathbb{R}^p	3
2.1	Représentation dans un espace euclidien	3
2.1.1	Représentation d'un individu par un point	3
2.1.2	Définition d'une distance entre les points A_i	4
2.1.3	Visualisation	6
2.2	Détermination du sous-espace de projection	6
2.2.1	Critère de détermination du sous-espace de projection	6
2.2.2	Détermination des axes principaux	8
2.2.3	Qualité globale de représentation du nuage par projection	10
2.3	Projection du nuage	11
2.3.1	Projection des points A_i	11
2.3.2	Projection des axes de la base canonique de \mathbb{R}^p	12
3	Etude statistique dans le dual \mathbb{R}^{p*} de \mathbb{R}^p	12
3.1	Définition des facteurs principaux	12
3.2	Interprétation statistique de l'ACP	13
3.3	Deuxième interprétation statistique de l'ACP	14
4	Etude dans l'espace des caractères \mathbb{R}^n	16
4.1	Structure euclidienne de \mathbb{R}^n	16
4.2	Les composantes principales	17
4.3	Le schéma de dualité	19
4.4	Analyse factorielle d'un tableau de distances	19
4.5	Le cercle des corrélations	20
5	Interprétation des résultats d'une ACP	21
5.1	Axes principaux à conserver	22
5.2	Etude des corrélations entre les caractères initiaux	22
5.3	Interprétation des facteurs	22
5.3.1	Analyse des corrélations des facteurs avec les caractères	23

5.3.2	Analyse de la position d'individus ayant une contribution importante	23
5.4	Interprétation de la position des projections des points-individus	24

1 Données et buts

Soit :

- $I = \{1, 2, \dots, n\}$ un ensemble d'individus;
- $P = \{p_1, p_2, \dots, p_n\}$ un ensemble de poids attribués aux individus, avec $0 < p_i < 1$, $\sum_{i=1}^n p_i = 1$;
- x^1, x^2, \dots, x^p des caractères quantitatifs mesurés sur les individus.

On centre les caractères x^j (on calcule les moyennes $\bar{x}^j = \sum_{i=1}^n p_i x_i^j$ et on centre les données) ; désormais, on note x_i^j la valeur centrée du caractère x^j pour l'individu i . On a le tableau de données suivant dans lequel les individus sont en ligne et les caractères en colonne :

$$X = \begin{matrix} & 1 & & & & & \\ & \vdots & & & & & \\ i & \cdots & \cdots & x_i^j & \cdots & \cdots & \\ & \vdots & & \vdots & & & \\ n & & & \vdots & & & \end{matrix} ;$$

On note $\underline{x}_i = \begin{pmatrix} x_i^1 \\ \vdots \\ \vdots \\ \vdots \\ x_i^p \end{pmatrix}$, $\underline{x}^j = \begin{pmatrix} x_1^j \\ \vdots \\ \vdots \\ \vdots \\ x_n^j \end{pmatrix}$. \underline{x}_i représente l'individu i et \underline{x}^j le caractère x^j .

Alors, $X = \begin{pmatrix} \underline{x}_1 \\ \vdots \\ \underline{x}'_i \\ \vdots \\ \underline{x}'_n \end{pmatrix} = (\underline{x}^1 \cdots \underline{x}^j \cdots \underline{x}^p)$.

On note D la matrice diagonale des poids des individus :

$$D = \begin{pmatrix} p_1 & & & \\ & \ddots & & \\ & & \ddots & 0 \\ 0 & & & p_2 \end{pmatrix}.$$

Le but de cette étude est de *décrire* (mise en évidence et analyse des ressemblances et dissemblances entre individus, des corrélations entre caractères) et *synthétiser* (détermination de facteurs combinaisons des caractères) le tableau de données X et d'en effectuer une *représentation graphique*. Elle est divisée en trois parties:

- 1) une étude dans l'espace des individus \mathbb{R}^p , dans laquelle on détermine les *axes principaux* ;
- 2) une étude dans le dual \mathbb{R}^{p*} de \mathbb{R}^p , dans laquelle on détermine les *facteurs principaux* ;
- 3) une étude dans l'espace des caractères \mathbb{R}^n , dans laquelle on détermine les *composantes principales*.

Cette analyse peut aussi conduire à remplacer un nombre important de caractères par un nombre réduit de facteurs non corrélés. En ce sens, elle peut être mise en oeuvre préalablement à une autre analyse statistique ; par exemple, avant d'effectuer une régression.

2 Etude géométrique dans l'espace des individus \mathbb{R}^p

Le but de cette étude est de déterminer une représentation graphique des individus par des points de telle manière que deux individus ayant des valeurs voisines (respectivement éloignées) pour les caractères soient représentés par des points voisins (respectivement éloignés).

2.1 Représentation dans un espace euclidien

2.1.1 Représentation d'un individu par un point

On représente l'individu i par le point A_i de R^p de coordonnées x_i^1, \dots, x_i^p :

$$i \mapsto A_i(\underline{x}_i).$$

On affecte le poids p_i au point A_i .

Le barycentre G des points (A_i, p_i) est à l'origine car les caractères sont centrés :

$$\overrightarrow{OG} = \sum_{i=1}^n p_i \overrightarrow{OA_i} \text{ a pour } j^{\text{ième}} \text{ composante } g^j = \sum_{i=1}^n p_i x_i^j = 0.$$

2.1.2 Définition d'une distance entre les points A_i

On convient de représenter la différence entre deux individus vis-à-vis des caractères par un nombre qui est une distance euclidienne entre les points représentatifs des individus.

On choisit une métrique M dans \mathbb{R}^p à laquelle est associée la distance euclidienne d . Pour deux points $A_i(\underline{x}_i)$ et $A_{i'}(\underline{x}_{i'})$, on a :

$$d^2(A_i, A_{i'}) = \|\underline{x}_i - \underline{x}_{i'}\|^2 = (\underline{x}_i - \underline{x}_{i'})' M (\underline{x}_i - \underline{x}_{i'}).$$

Différents choix de métriques sont possibles.

1) Distance euclidienne usuelle

On prend $M = I$.

$$d(A_i, A_{i'}) = \sqrt{\sum_{j=1}^p (x_i^j - x_{i'}^j)^2}$$

Cette distance présente un inconvénient lorsque les caractères mesurés sont hétérogènes, c'est-à-dire ne s'expriment pas dans la même unité, comme le montre l'exemple suivant.

Exemple On a mesuré dans deux pays les exportations, en millions de dollars, et le taux d'escompte, en pourcentage:

	Exportations	Taux d'escompte
Pays 1	28 054	7,00
Pays 2	2 306	5,50
Différence	25 748	1,50

On constate que le taux d'escompte n'interviendra pratiquement pas dans la distance entre les deux pays. De plus, les différences calculées dépendent des unités choisies. ■

2) Distance de l'ACP normée

On note s^j l'écart-type du caractère x^j : $s^j = \sqrt{\sum_{i=1}^n p_i (x_i^j)^2}$. On définit la distance:

$$d(A_i, A_{i'}) = \sqrt{\sum_{j=1}^p \left(\frac{x_i^j - x_{i'}^j}{s^j} \right)^2} = \sqrt{\sum_{j=1}^p \frac{1}{(s^j)^2} (x_i^j - x_{i'}^j)^2}$$

Les différences entre les valeurs de deux caractères s'expriment alors en écarts-types.

- Elles ne dépendent plus de l'unité de mesure choisie, car l'unité de s^j est la même que celle de x_i^j .
- Elles sont du même ordre de grandeur, comprises en général entre 0 et 6.

En effet, le caractère x^j étant centré, d'après l'inégalité de Bienaymé-Tchébychev, au moins $\frac{8}{9}$ de ses valeurs x_i^j sont comprises entre $-3s^j$ et $+3s^j$.

Choisir cette distance revient à prendre dans \mathbb{R}^p la métrique diagonale suivante, appelée métrique de l'inverse des variances :

$$M = \begin{pmatrix} \frac{1}{(s^1)^2} & & & & \\ & \ddots & & & \\ & & \frac{1}{(s^j)^2} & & \\ & & & \ddots & \\ & & & & \frac{1}{(s^p)^2} \end{pmatrix}.$$

On peut aussi considérer que ceci revient à prendre les données centrées réduites $\frac{x_i^j}{s^j}$ et à utiliser la métrique-identité.

3) Autres métriques possibles

On peut utiliser une métrique diagonale de poids attribués aux caractères:

$$M = \begin{pmatrix} q^1 & & & & \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & q^p \end{pmatrix}, \quad d(A_i, A_{i'}) = \sqrt{\sum_{j=1}^p q^j (x_i^j - x_{i'}^j)^2}.$$

C'est ce que l'on peut faire par exemple dans l'analyse factorielle multiple, qui est étudiée plus loin.

En analyse canonique généralisée, également étudiée plus loin, on utilise une métrique M diagonale par blocs:

$$M = \begin{pmatrix} M^1 & & & & \\ & \ddots & & & \\ & & M^2 & & 0 \\ 0 & & & \ddots & \\ & & & & M^q \end{pmatrix}.$$

2.1.3 Visualisation

Si l'on pouvait voir dans \mathbb{R}^p , le but de l'étude serait atteint. *On va alors projeter le nuage de points A_i sur un sous-espace de faible dimension, choisi de façon optimale.*

2.2 Détermination du sous-espace de projection

2.2.1 Critère de détermination du sous-espace de projection

On détermine un sous-espace de projection F_r , de dimension fixée r , tel que la projection du nuage des points A_i sur ce sous-espace soit l'image la plus fidèle possible du nuage initial.

1) Première expression

On note P_i la projection de A_i sur F_r . On a:

$$d(P_i, P_{i'}) = \|\overrightarrow{P_i P_{i'}}\| \leq \|\overrightarrow{A_i A_{i'}}\| = d(A_i, A_{i'}).$$

Donc:

$$\underbrace{\sum_{i=1}^n \sum_{i'=1}^n p_i p_{i'} \|\overrightarrow{P_i P_{i'}}\|^2}_{\text{dépend du sous-espace } F_r} \leq \underbrace{\sum_{i=1}^n \sum_{i'=1}^n p_i p_{i'} \|\overrightarrow{A_i A_{i'}}\|^2}_{\text{nombre fixe}}$$

Pour que, dans leur ensemble, les distances entre les points P_i soient les plus proches possibles des distances entre les points A_i , on détermine F_r de telle manière que la somme

$$\sum_{i=1}^n \sum_{i'=1}^n p_i p_{i'} \|\overrightarrow{P_i P_{i'}}\|^2 \text{ soit maximale.}$$

Critère F_r est un sous-espace de dimension r qui rend maximale la somme $\sum_{i=1}^n \sum_{i'=1}^n p_i p_{i'} \|\overrightarrow{P_i P_{i'}}\|^2$.

2) Deuxième expression

On note G' la projection du barycentre G sur F_r : G' est le barycentre des points (P_i, p_i) .

Lemme $\sum_{i=1}^n \sum_{i'=1}^n p_i p_{i'} \|\overrightarrow{P_i P_{i'}}\|^2 = 2 \sum_{i=1}^n p_i \|\overrightarrow{G' P_i}\|^2$.

Démonstration

$$\|\overrightarrow{P_i P_{i'}}\|^2 = \|\overrightarrow{P_i G'} + \overrightarrow{G' P_{i'}}\|^2 = \|\overrightarrow{P_i G'}\|^2 + \|\overrightarrow{G' P_{i'}}\|^2 + 2 \langle \overrightarrow{P_i G'}, \overrightarrow{G' P_{i'}} \rangle$$

$$\begin{aligned} \sum_{i=1}^n \sum_{i'=1}^n p_i p_{i'} \|\overrightarrow{P_i P_{i'}}\|^2 &= \sum_{i=1}^n \sum_{i'=1}^n p_i p_{i'} (\|\overrightarrow{P_i G'}\|^2 + \|\overrightarrow{G' P_{i'}}\|^2 + 2 \langle \overrightarrow{P_i G'}, \overrightarrow{G' P_{i'}} \rangle) \\ &= 2 \sum_{i=1}^n p_i \|\overrightarrow{P_i G'}\|^2 + 2 \langle \sum_{i=1}^n p_i \overrightarrow{P_i G'}, \sum_{i'=1}^n p_{i'} \overrightarrow{G' P_{i'}} \rangle \\ &= 2 \sum_{i=1}^n p_i \|\overrightarrow{G' P_i}\|^2 \quad \text{car} \quad \sum_{i'=1}^n p_{i'} \overrightarrow{G' P_{i'}} = \vec{0}. \blacksquare \end{aligned}$$

Dans le triangle $G' A_i P_i$ rectangle en P_i , on a :

$$\|\overrightarrow{G' A_i}\|^2 = \|\overrightarrow{G' P_i}\|^2 + \|\overrightarrow{A_i P_i}\|^2.$$

$$\sum_{i=1}^n p_i \|\overrightarrow{A_i P_i}\|^2 = \sum_{i=1}^n p_i \|\overrightarrow{G' A_i}\|^2 - \underbrace{\sum_{i=1}^n p_i \|\overrightarrow{G' P_i}\|^2}_{\max}$$

D'après le lemme, le critère précédent revient à maximiser $\sum_{i=1}^n p_i \|\overrightarrow{G' P_i}\|^2$.

$$\begin{aligned} \sum_{i=1}^n p_i \|\overrightarrow{G' A_i}\|^2 &= \sum_{i=1}^n p_i \|\overrightarrow{G' G} + \overrightarrow{G A_i}\|^2 \\ &= \sum_{i=1}^n p_i (\|\overrightarrow{G' G}\|^2 + 2 \langle \overrightarrow{G' G}, \overrightarrow{G A_i} \rangle + \|\overrightarrow{G A_i}\|^2) \\ &= \|\overrightarrow{G' G}\|^2 + 2 \langle \overrightarrow{G' G}, \sum_{i=1}^n p_i \overrightarrow{G A_i} \rangle + \sum_{i=1}^n p_i \|\overrightarrow{G A_i}\|^2 \end{aligned}$$

Or: $\sum_{i=1}^n p_i \overrightarrow{G A_i} = \vec{0}$. Donc:

$$\sum_{i=1}^n p_i \|\overrightarrow{A_i P_i}\|^2 = \|\overrightarrow{G' G}\|^2 + \sum_{i=1}^n p_i \|\overrightarrow{G A_i}\|^2 - \sum_{i=1}^n p_i \|\overrightarrow{G' P_i}\|^2.$$

$$\sum_{i=1}^n p_i \|\overrightarrow{G'P_i}\|^2 \max \quad \text{et} \quad G' = G \iff \sum_{i=1}^n p_i \|A_i P_i\|^2 \min$$

C'est ce critère des moindres carrés que l'on utilise pour définir le sous-espace affine F_r , qui contient alors le barycentre G .

Critère F_r est un sous-espace affine de dimension r qui rend minimale la somme $\sum_{i=1}^n p_i \|\overrightarrow{A_i P_i}\|^2$: il ajuste au mieux le nuage des points (A_i, p_i) au sens des moindres carrés, les distances des points A_i à F_r étant prises orthogonalement.

On appelle :

- inertie des points A_i par rapport à G : $\sum_{i=1}^n p_i \|\overrightarrow{GA_i}\|^2$;
- inertie des points P_i par rapport à G : $\sum_{i=1}^n p_i \|\overrightarrow{GP_i}\|^2$;
- inertie des points A_i par rapport à F_r : $\sum_{i=1}^n p_i \|A_i P_i\|^2$.

2.2.2 Détermination des axes principaux

On note $(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_r)$ une base M -orthonormée de F_r .

Lemme
$$\sum_{i=1}^n p_i \|\overrightarrow{GP_i}\|^2 = \sum_{k=1}^r \underline{u}'_k M X' D X M \underline{u}_k.$$

Démonstration

1) Notons P_i^k la projection de P_i sur l'axe (G, \vec{u}_k) . D'après le théorème des trois perpendiculaires, P_i^k est la projection de A_i sur l'axe (G, \vec{u}_k) .

En effet, comme \vec{u}_k est unitaire et $\overrightarrow{A_i P_i}$ orthogonal à \vec{u}_k :

$$\overrightarrow{GP_i^k} = \langle \overrightarrow{GP_i}, \vec{u}_k \rangle = \langle \overrightarrow{GA_i} + \overrightarrow{A_i P_i}, \vec{u}_k \rangle = \langle \overrightarrow{GA_i}, \vec{u}_k \rangle.$$

$\overrightarrow{GP_i^k} = \langle \overrightarrow{GA_i}, \vec{u}_k \rangle$; donc, P_i^k est la projection orthogonale de A_i sur (G, \vec{u}_k) .

2) Comme la base $(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_r)$ est orthonormée, on a :

$$\|\overrightarrow{GP_i}\|^2 = \sum_{k=1}^r \|\overrightarrow{GP_i^k}\|^2 \implies \sum_{i=1}^n p_i \|\overrightarrow{GP_i}\|^2 = \sum_{k=1}^r \sum_{i=1}^n p_i \|\overrightarrow{GP_i^k}\|^2.$$

3) $\overrightarrow{GP_i^k} = \overrightarrow{GA_i} \cdot \vec{u}_k = \underline{x}'_i M \underline{u}_k = \underline{u}'_k M \underline{x}_i$

$$\sum_{i=1}^n p_i \|\overrightarrow{GP_i^k}\|^2 = \sum_{i=1}^n p_i \underline{u}'_k M \underline{x}_i \underline{x}'_i M \underline{u}_k = \underline{u}'_k M \left(\sum_{i=1}^n p_i \underline{x}_i \underline{x}'_i \right) M \underline{u}_k$$

$$\sum_{i=1}^n p_i \underline{x}_i \underline{x}'_i = (\underline{x}_1 \cdots \underline{x}_i \cdots \underline{x}_n) \begin{pmatrix} p_1 & & & \\ & \ddots & & \\ & & \ddots & \\ 0 & & & p_n \end{pmatrix} \begin{pmatrix} \underline{x}'_1 \\ \vdots \\ \underline{x}'_i \\ \vdots \\ \underline{x}'_n \end{pmatrix} = X'DX$$

L'élément (l, m) de $X'DX$ est égal à $\sum_{i=1}^n p_i x_i^l x_i^m$. On a :

$$\begin{aligned} \sum_{i=1}^n p_i \|\overrightarrow{GP_i^k}\|^2 &= \underline{u}'_k M X' D X M \underline{u}_k \\ \sum_{i=1}^n p_i \|\overrightarrow{GP_i}\|^2 &= \sum_{k=1}^r \underline{u}'_k M X' D X M \underline{u}_k. \blacksquare \end{aligned}$$

Le problème est donc maintenant de déterminer une base $(\underline{u}_1, \dots, \underline{u}_r)$ M -orthonormée telle que $\sum_{k=1}^r \underline{u}'_k M X' D X M \underline{u}_k$ soit maximale.

Théorème *A partir d'un sous-espace F_r vérifiant le critère, on peut construire un sous-espace F_{r+1} le vérifiant en lui adjoignant une droite M -orthogonale $\Delta_{\underline{u}_{r+1}}$ telle que la forme quadratique $\underline{u}'_{r+1} M X' D X M \underline{u}_{r+1}$ soit maximale sous la contrainte $\underline{u}'_{r+1} M \underline{u}_{r+1} = 1$.*

Ceci revient à dire : pour maximiser $\sum_{k=1}^{r+1} \underline{u}'_k M X' D X M \underline{u}_k$ sous la contrainte $(\vec{u}_1, \dots, \vec{u}_r, \vec{u}_{r+1})$

M -orthonormée, on maximise $\sum_{k=1}^r \underline{u}'_k M X' D X M \underline{u}_k$ sous la contrainte $(\vec{u}_1, \dots, \vec{u}_r)$ M -orthonormée et $\underline{u}'_{r+1} M X' D X M \underline{u}_{r+1}$ sous les contraintes \vec{u}_{r+1} unitaire et \vec{u}_{r+1} orthogonal à \vec{u}_j pour $j = 1, \dots, r$.

En pratique, on peut donc procéder pas à pas pour déterminer les axes principaux :

$$\left\{ \begin{array}{l} F_1 = (G, \vec{u}_1) \quad : \underline{u}'_1 M X' D X M \underline{u}_1 \text{ max, } \underline{u}'_1 M \underline{u}_1 = 1. \\ F_2 = (G, \vec{u}_1, \vec{u}_2) \quad : \underline{u}'_2 M X' D X M \underline{u}_2 \text{ max, } \begin{array}{l} \underline{u}'_2 M \underline{u}_2 = 1, \\ \underline{u}'_2 M \underline{u}_1 = 0. \end{array} \\ F_3 = (G, \vec{u}_1, \vec{u}_2, \vec{u}_3) \quad : \underline{u}'_3 M X' D X M \underline{u}_3 \text{ max, } \begin{array}{l} \underline{u}'_3 M \underline{u}_3 = 1, \\ \underline{u}'_3 M \underline{u}_1 = 0, \\ \underline{u}'_3 M \underline{u}_2 = 0. \end{array} \\ \vdots \end{array} \right.$$

Théorème \vec{u}_{r+1} est vecteur propre de $X'DXM$ associé à sa $(r+1)^{\text{ième}}$ plus grande valeur propre λ_{r+1} . On a : $\sum_{i=1}^n p_i \|\overrightarrow{GP_i^{r+1}}\|^2 = \lambda_{r+1}$.

Démonstration

\vec{u}_{r+1} rend maximale $\underline{u}'_{r+1}MX'DXM\underline{u}_{r+1}$
 sous les contraintes : $\begin{cases} \underline{u}'_{r+1}M\underline{u}_{r+1} = 1, \\ \underline{u}'_{r+1}M\underline{u}_j = 0, j = 1, \dots, r. \end{cases}$

La matrice $A = MX'DXM$ est symétrique. D'après le théorème de maximisation sous contraintes d'une forme quadratique étudié dans le chapitre précédent, \vec{u}_{r+1} est vecteur propre de $M^{-1}A = M^{-1}MX'DXM = X'DXM$ associé à la $(r + 1)^{i\grave{e}me}$ plus grande valeur propre. D'après la démonstration du lemme, on a :

$$\sum_{i=1}^n p_i \|\overrightarrow{GP_i^{r+1}}\|^2 = \underline{u}'_{r+1}MX'DXM\underline{u}_{r+1} = \lambda_{r+1}\underline{u}'_{r+1}M\underline{u}_{r+1} = \lambda_{r+1}. \blacksquare$$

Définition L 'axe (G, \vec{u}_k) est appelé le $k^{i\grave{e}me}$ axe principal.

Remarque $V = X'DX$ est la matrice des covariances des caractères x^1, x^2, \dots, x^p . Si les caractères sont réduits, $X'DX$ est la matrice de corrélation (matrice des coefficients de corrélation linéaire) dont l'élément général (l, m) est :

$$\sum_{i=1}^n p_i \frac{x_i^l}{s^l} \frac{x_i^m}{s^m} = \frac{1}{s^l s^m} \sum_{i=1}^n p_i x_i^l x_i^m = r(x^l, x^m). \blacksquare$$

2.2.3 Qualité globale de représentation du nuage par projection

Définition On appelle qualité globale de représentation du nuage des points A_i par projection sur F_r le rapport

$$\frac{\sum_{i=1}^n \sum_{i'=1}^n p_i p_{i'} \|\overrightarrow{P_i P_{i'}}\|^2}{\sum_{i=1}^n \sum_{i'=1}^n p_i p_{i'} \|\overrightarrow{A_i A_{i'}}\|^2} = \frac{\sum_{i=1}^n p_i \|\overrightarrow{GP_i}\|^2}{\sum_{i=1}^n p_i \|\overrightarrow{GA_i}\|^2} = \frac{\text{inertie des projections } P_i \text{ par rapport à } G}{\text{inertie des projections } A_i \text{ par rapport à } G}.$$

On l'appelle aussi pourcentage d'inertie expliquée par le sous-espace F_r .

Proposition Elle est égale à $\frac{\lambda_1 + \lambda_2 + \dots + \lambda_r}{\lambda_1 + \lambda_2 + \dots + \lambda_p} = \frac{\lambda_1 + \lambda_2 + \dots + \lambda_r}{\text{Trace}(X'DXM)}$. C'est la somme des qualités globales de représentation par projection sur les axes principaux qui engendrent F_r .

Démonstration

$$\sum_{i=1}^n p_i \|\overrightarrow{GP_i}\|^2 = \sum_{k=1}^r \sum_{i=1}^n p_i \|\overrightarrow{GP_i^k}\|^2 = \sum_{k=1}^r \lambda_k$$

Comme $(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p)$ est une base M -orthonormé de \mathbb{R}^p , on a : $\|\vec{GA}_i\|^2 = \sum_{k=1}^p \|\vec{GP}_i^k\|^2$.
Donc:

$$\sum_{i=1}^n p_i \|\vec{GA}_i\|^2 = \sum_{k=1}^p \sum_{i=1}^n p_i \|\vec{GP}_i^k\|^2 = \sum_{k=1}^p \lambda_k.$$

La qualité globale de représentation par projection sur l'axe principal (G, \vec{u}_k) est égale à :

$$\frac{\sum_{i=1}^n p_i \|\vec{GP}_i^k\|^2}{\sum_{i=1}^n p_i \|\vec{GA}_i\|^2} = \frac{\lambda_k}{\lambda_1 + \dots + \lambda_p}. \blacksquare$$

Remarque Dans le cas de l'ACP normée, la matrice $X'DXM$ a des 1 sur la diagonale. Alors $\text{Trace}(X'DXM) = p$. \blacksquare

2.3 Projection du nuage

2.3.1 Projection des points A_i

Proposition L'abscisse de la projection du point A_i sur l'axe principal (G, \vec{u}_k) , \vec{GP}_i^k , est égale à $\underline{x}'_i M \underline{u}_k$.

Définition On appelle qualité individuelle de représentation du point A_i par projection sur l'axe principal (G, \vec{u}_k) le rapport $\frac{\|\vec{GP}_i^k\|^2}{\|\vec{GA}_i\|^2} = \cos^2(\vec{GA}_i, \vec{GP}_i^k)$. On appelle qualité individuelle de représentation de A_i par projection sur le sous-espace F_r le rapport $\frac{\|\vec{GP}_i\|^2}{\|\vec{GA}_i\|^2} = \cos^2(\vec{GA}_i, \vec{GP}_i)$.

On a :

$$\begin{aligned} \|\vec{GP}_i\|^2 &= \sum_{k=1}^r \|\vec{GP}_i^k\|^2; \quad \frac{\|\vec{GP}_i\|^2}{\|\vec{GA}_i\|^2} = \sum_{k=1}^r \frac{\|\vec{GP}_i^k\|^2}{\|\vec{GA}_i\|^2}; \\ \|\vec{GP}_i^k\|^2 &= (\underline{x}'_i M \underline{u}_k)^2; \quad \|\vec{GA}_i\|^2 = \underline{x}'_i M \underline{x}_i. \end{aligned}$$

Proposition La qualité individuelle de représentation de A_i par projection sur F_r est égale à $\sum_{k=1}^r \frac{(\underline{x}'_i M \underline{u}_k)^2}{\underline{x}'_i M \underline{x}_i}$; c'est la somme des qualités individuelles de représentation par projection sur les axes principaux qui engendrent F_r .

Interprétation On doit interpréter avec prudence la position relative par rapport aux autres de la projection d'un point ayant une faible qualité de représentation. La proximité

entre deux points ayant une faible qualité de représentation n'implique pas que ces individus soient semblables vis-à-vis de tous les caractères.

Définition On appelle contribution de la projection P_i^k du point A_i sur l'axe principal (G, \underline{u}_k) à l'inertie expliquée par (G, \underline{u}_k) le rapport $\frac{p_i \|\overrightarrow{GP_i^k}\|^2}{\sum_{l=1}^n p_l \|\overrightarrow{GP_l^k}\|^2}$.

On a : $\sum_{l=1}^n p_l \|\overrightarrow{GP_l^k}\|^2 = \lambda_k$.

Proposition Cette contribution est égale à $\frac{p_i (\underline{x}_i' M \underline{u}_k)^2}{\lambda_k}$.

Interprétation L'examen des contributions permet de repérer d'éventuels individus aberrants ayant une contribution à l'inertie expliquée par un axe principal très forte par rapport à celle des autres, et qui peuvent parfois expliquer à eux seuls la construction de cet axe. Dans certains cas, on peut être amené à éliminer ces individus et à les faire figurer en individus supplémentaires ou passifs.

2.3.2 Projection des axes de la base canonique de \mathbb{R}^p

Soit $(\vec{e}_1, \dots, \vec{e}_p)$ la base canonique de \mathbb{R}^p . L'axe (G, \vec{e}_j) représente le $j^{\text{ième}}$ caractère x^j .

On note u_k^l la $l^{\text{ième}}$ composante de \underline{u}_k .

Proposition La mesure algébrique de la projection de \vec{e}_j sur l'axe principal (G, \vec{u}_k) est égale à $\langle \vec{e}_j, \vec{u}_k \rangle = e_j' M \underline{u}_k = \sum_{l=1}^p M_{jl} u_k^l$. Pour M diagonale, elle est égale à $M_{jj} u_k^j$.

3 Etude statistique dans le dual \mathbb{R}^{p*} de \mathbb{R}^p

Dans l'étude précédente, on a déterminé des axes principaux (G, \vec{u}_k) dans \mathbb{R}^p .

Dans \mathbb{R}^p , un axe de la base canonique représente un caractère ; un axe principal représente un "nouveau caractère", combinaison linéaire des caractères, que l'on appelle facteur principal. L'introduction de cette notion va permettre d'interpréter statistiquement l'étude effectuée.

3.1 Définition des facteurs principaux

Définition Le $k^{\text{ième}}$ facteur principal est une combinaison linéaire des caractères x^1, x^2, \dots, x^p centrés telle que sa valeur pour l'individu i soit l'abscisse de la projection P_i^k du point A_i sur l'axe principal (G, \vec{u}_k) , $\overrightarrow{GP_i^k} = \underline{x}_i' M \underline{u}_k$.

Notons $\underline{a}_k = M\underline{u}_k = \begin{pmatrix} a_k^1 \\ \vdots \\ a_k^p \end{pmatrix}$.

La valeur du $k^{\text{ième}}$ facteur par l'individu i est par définition :

$$\underline{x}'_i M\underline{u}_k = \underline{x}'_i \underline{a}_k = (x_i^1 \dots x_i^p) \begin{pmatrix} a_k^1 \\ \vdots \\ a_k^p \end{pmatrix} = a_k^1 x_i^1 + \dots + a_k^p x_i^p.$$

Le $k^{\text{ième}}$ facteur est la combinaison linéaire $a_k^1 x^1 + a_k^2 x^2 + \dots + a_k^p x^p$ des caractères x^1, \dots, x^p centrés.

Définition Le vecteur $\underline{a}_k = M\underline{u}_k$ des coefficients de la combinaison linéaire qui définit le $k^{\text{ième}}$ facteur est un élément du dual \mathbb{R}^{p*} de \mathbb{R}^p que l'on appelle aussi $k^{\text{ième}}$ facteur principal.

On définit l'application M de \mathbb{R}^p dans \mathbb{R}^{p*} :

$$\begin{aligned} M &: \mathbb{R}^p \longrightarrow \mathbb{R}^{p*} \\ \underline{u}_k &\longmapsto \underline{a}_k = M\underline{u}_k \end{aligned}$$

Théorème \underline{a}_k est vecteur propre M^{-1} -unitaire de la matrice $MX'DX$ associé à sa $k^{\text{ième}}$ plus grande valeur propre λ_k .

Démonstration

\underline{u}_k est vecteur propre M -unitaire de $X'DXM$ associé à sa $k^{\text{ième}}$ plus grande valeur propre λ_k : $X'DXM \underline{u}_k = \lambda_k \underline{u}_k$.

$$\begin{aligned} X'DXM \underline{u}_k &= \lambda_k \underline{u}_k \iff MX'DXM \underline{u}_k = \lambda_k M\underline{u}_k \iff MX'DX \underline{a}_k = \lambda_k \underline{a}_k \\ \underline{u}'_k M \underline{u}_k &= 1 \iff \underline{u}'_k M M^{-1} M \underline{u}_k = 1 \iff \underline{a}'_k M^{-1} \underline{a}_k = 1. \blacksquare \end{aligned}$$

3.2 Interprétation statistique de l'ACP

Théorème Pour $\lambda_{k-1} \neq 0$, \underline{a}_k est solution du problème de la recherche de $\underline{a} \in \mathbb{R}^{p*}$ tel que $\underline{a}' X' D X \underline{a}$ soit maximal sous les contraintes $\underline{a}' X' D X \underline{a}_j = 0$, $j = 1, 2, \dots, k-1$, $\underline{a}' M^{-1} \underline{a} = 1$.

Démonstration

\underline{u}_k est solution du problème de la recherche de $\underline{u} \in \mathbb{R}^p$ tel que $\underline{u}' M X' D X M \underline{u}$ soit maximal sous les contraintes $\underline{u}' M \underline{u}_j = 0$, $j = 1, \dots, k-1$ et $\underline{u}' M \underline{u} = 1$.

$$\begin{aligned} \underline{u}' M X' D X M \underline{u} \text{ maximal} &\iff \underline{a}' X' D X \underline{a} \text{ maximal} ; \\ \underline{u}' M \underline{u} = 1 &\iff \underline{u}' M M^{-1} M \underline{u} = 1 \iff \underline{a}' M^{-1} \underline{a} = 1 ; \end{aligned}$$

\underline{u}_j est vecteur propre de $X'DXM$: $X'DXM\underline{u}_j = \lambda_j\underline{u}_j \iff \underline{u}_j = \frac{1}{\lambda_j}X'DXM\underline{u}_j$, pour $\lambda_j \neq 0$; donc :
 $\underline{u}'M\underline{u}_j = 0 \iff \underline{u}'MX'DXM\underline{u}_j = 0 \iff \underline{a}'X'DX\underline{a}_j = 0$. ■

Interprétation

\underline{a}_k est le vecteur des coefficients d'une combinaison linéaire des caractères x^1, x^2, \dots, x^p .

$$X\underline{a}_k = \begin{pmatrix} x_i^1 & \dots & x_i^p \end{pmatrix} \begin{pmatrix} a_k^1 \\ \vdots \\ a_k^p \end{pmatrix} = \begin{pmatrix} \vdots \\ \sum_{j=1}^p a_k^j x_i^j \\ \vdots \end{pmatrix} \leftarrow \text{valeur du } k^{\text{ième}} \text{ facteur pour}$$

l'individu i .

$$\underline{a}_k'X'DX\underline{a}_k = (X\underline{a}_k)'D(X\underline{a}_k) = \begin{pmatrix} \dots & \sum_{j=1}^p a_k^j x_i^j & \dots \end{pmatrix} \begin{pmatrix} p_1 & & & 0 \\ & \ddots & & \\ 0 & & \ddots & \\ & & & p_n \end{pmatrix} \begin{pmatrix} \vdots \\ \sum_{j=1}^p a_k^j x_i^j \\ \vdots \end{pmatrix} = \sum_{i=1}^n p_i \left(\sum_{j=1}^p a_k^j x_i^j \right)^2 : \text{c'est la vari-}$$

ance des valeurs du $k^{\text{ième}}$ facteur.

$\underline{a}_k'X'DX\underline{a}_j$ est la covariance du $k^{\text{ième}}$ et du $j^{\text{ième}}$ facteur.

Proposition *Le $k^{\text{ième}}$ facteur de l'ACP est une combinaison linéaire des caractères x^1, x^2, \dots, x^p centrés de variance maximale sous les contraintes qu'elle soit non corrélée aux facteurs précédents et la contrainte de normalisation $\underline{a}_k'M^{-1}\underline{a}_k = 1$.*

On peut présenter l'ACP comme la recherche d'une combinaison linéaire des caractères centrés de variance maximale, puis d'une deuxième combinaison non corrélée à la première et de variance maximale, puis d'une troisième combinaison non corrélée aux deux premières et de variance maximale et ainsi de suite. L'ACP conduit à remplacer p caractères par un nombre réduit de facteurs non corrélés qui peuvent alors être utilisés dans une autre analyse.

Remarque La variance du $k^{\text{ième}}$ facteur est $\underline{a}_k'X'DX\underline{a}_k = \underline{u}_k'MX'DXM\underline{u}_k = \sum_{i=1}^r p_i \left\| \overrightarrow{GP_i^k} \right\| = \lambda_k$; la qualité globale de représentation par projection sur le $k^{\text{ième}}$ axe principal, $\frac{\lambda_k}{\lambda_1 + \dots + \lambda_p}$, est aussi appelée *pourcentage de variance expliquée par le $k^{\text{ième}}$ facteur*. ■

3.3 Deuxième interprétation statistique de l'ACP

On note $V = X'DX$ la matrice des covariances de x^1, \dots, x^p . On la suppose définie positive (ceci est réalisé si et seulement si il n'existe pas de relation linéaire entre les caractères x^1, \dots, x^p centrés).

Théorème Le vecteur $\underline{b}_k = \frac{\underline{a}_k}{\sqrt{\lambda_k}}$ de \mathbb{R}^{p*} est solution du problème de la recherche de $\underline{b} \in \mathbb{R}^{p*}$ tel que $\underline{b}'VMV\underline{b}$ soit maximal sous les contraintes $\underline{b}'V\underline{b}_j = 0, j = 1, \dots, k-1, \underline{b}'V\underline{b} = 1$.

Démonstration

Soit $A = VMV$. On cherche \underline{b} tel que $\underline{b}'A\underline{b}$ soit maximal sous les contraintes $\underline{b}'V\underline{b}_j = 0, j = 1, \dots, k-1$ et $\underline{b}'V\underline{b} = 1$. Toute solution de ce problème est vecteur propre de $V^{-1}A = MV$ associé à la $k^{\text{ième}}$ plus grande valeur propre, comme \underline{a}_k .

Mais, $\underline{a}'_k V \underline{a}_k = \underline{a}'_k X' DX \underline{a}_k = \lambda_k$ (variance du $k^{\text{ième}}$ facteur). On peut prendre $\underline{b}_k = \frac{\underline{a}_k}{\sqrt{\lambda_k}}$. ■

Cas où la métrique M est diagonale

On note M_{jj} le $j^{\text{ième}}$ élément diagonal de M et \underline{e}_j le $j^{\text{ième}}$ vecteur de la base canonique de \mathbb{R}^p .

$$\underline{e}_j \underline{e}'_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} (0 \dots 1 \dots 0) = \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \ddots \\ 0 & & & & 0 \end{pmatrix} \leftarrow j$$

$$M = \begin{pmatrix} M_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \ddots & \\ & & & & M_{pp} \end{pmatrix} = \sum_{j=1}^p M_{jj} \underline{e}_j \underline{e}'_j$$

$$\underline{b}'_k V M V \underline{b}_k = \underline{b}'_k V \sum_{j=1}^p M_{jj} \underline{e}_j \underline{e}'_j V \underline{b}_k$$

$$= \sum_{j=1}^p M_{jj} (\underline{b}'_k V \underline{e}_j)^2$$

$$\underline{b}'_k V \underline{e}_j = \frac{1}{\sqrt{\lambda_k}} \underline{a}'_k V \underline{e}_j = \frac{1}{\sqrt{\lambda_k}} \underline{a}'_k X' DX \underline{e}_j = \frac{1}{\sqrt{\lambda_k}} (X \underline{a}_k)' DX \underline{e}_j$$

$X \underline{a}_k$ est le vecteur des valeurs du $k^{\text{ième}}$ facteur pour les n individus;

$$X \underline{e}_j = (\underline{x}^1 \dots \underline{x}^j \dots \underline{x}^p) \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \underline{x}^j.$$

$(X\underline{a}_k)'D(X\underline{e}_j)$ est la covariance du $k^{\text{ième}}$ facteur et du caractère x^j ; $\frac{1}{\sqrt{\lambda_k}}(X\underline{a}_k)'D(X\underline{e}_j)$ est la covariance du $k^{\text{ième}}$ facteur réduit, noté v^k , et du caractère x^j , $s_{v^k x^j}$.

$$\text{Donc, } \underline{b}'_k V M V \underline{b}_k = \sum_{j=1}^p M_{jj} s_{v^k x^j}^2.$$

On en déduit la proposition suivante.

Proposition Dans le cas où la métrique M est diagonale, le $k^{\text{ième}}$ facteur réduit est une combinaison linéaire v^k des caractères x^1, x^2, \dots, x^p centrés qui rend maximale $\sum_{j=1}^p M_{jj} s_{v^k x^j}^2$ sous les contraintes d'être non corrélée aux facteurs précédents.

Cas de l'ACP normée

$$M_{jj} = \frac{1}{(s^j)^2} ; \sum_{j=1}^p M_{jj} s_{v^k x^j}^2 = \sum_{j=1}^p \frac{1}{(s^j)^2} s_{v^k x^j}^2 = \sum_{j=1}^p r_{v^k x^j}^2.$$

Proposition Le $k^{\text{ième}}$ facteur réduit de l'ACP normée est une combinaison linéaire v^k des caractères x^1, \dots, x^p centrés qui rend maximale $\sum_{j=1}^p r_{v^k x^j}^2$ sous les contraintes d'être non corrélée aux facteurs précédents.

4 Etude dans l'espace des caractères \mathbb{R}^n

4.1 Structure euclidienne de \mathbb{R}^n

Dans \mathbb{R}^n , on représente chaque caractère x^j par le point B^j de coordonnées (x_1^j, \dots, x_n^j) : $B^j(\underline{x}^j)$.

On munit \mathbb{R}^n de la métrique des points des individus:

$$D = \begin{pmatrix} p_1 & & & 0 \\ & \ddots & & \\ 0 & & \ddots & \\ & & & p_n \end{pmatrix}.$$

On a alors :

$$\begin{aligned} \cdot \langle \underline{x}^j, \underline{x}^k \rangle &= (\underline{x}^j)' D \underline{x}^k = (x_1^j, \dots, x_n^j) \begin{pmatrix} p_1 & & & 0 \\ & \ddots & & \\ 0 & & \ddots & \\ & & & p_n \end{pmatrix} \begin{pmatrix} x_1^k \\ \vdots \\ \vdots \\ x_n^k \end{pmatrix} \\ &= \sum_{i=1}^n p_i x_i^j x_i^k : \text{covariance de } x^j \text{ et de } x^k ; \end{aligned}$$

- $\|\underline{x}^j\|$: écart-type de \underline{x}^j ;
- $\cos(\underline{x}^j, \underline{x}^k) = \frac{\langle \underline{x}^j, \underline{x}^k \rangle}{\|\underline{x}^j\| \|\underline{x}^k\|}$: coefficient de corrélation linéaire de x^j et x^k .

Remarque Le caractère x^j est centré :

$$\sum_{i=1}^n p_i x_i^j = 0 = \langle \underline{x}^j, \underline{u} \rangle \quad \text{avec } \underline{u} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Le caractère x^j est centré si et seulement si \underline{x}^j appartient à l'orthogonal à la droite $\Delta \underline{u}$ engendrée par \underline{u} . ■

4.2 Les composantes principales

La valeur du $k^{\text{ième}}$ facteur pour l'individu i est, par définition, $\underline{x}'_i M \underline{u}_k = \underline{x}'_i \underline{a}_k$.

Définition Le vecteur \underline{c}^k des valeurs du $k^{\text{ième}}$ facteur pour les n individus est appelé la $k^{\text{ième}}$ composante principale.

Expression

$$\underline{c}^k = \begin{pmatrix} \underline{x}'_1 \underline{a}_k \\ \underline{x}'_2 \underline{a}_k \\ \vdots \\ \underline{x}'_n \underline{a}_k \end{pmatrix}_{(n, 1)} = \begin{pmatrix} \underline{x}'_1 \\ \underline{x}'_2 \\ \vdots \\ \underline{x}'_n \end{pmatrix}_{(n, p)} \underline{a}_k = X \underline{a}_k = X M \underline{u}_k.$$

On définit l'application :

$$X : \mathbb{R}^{p*} \rightarrow \mathbb{R}^n \\ \underline{a}_k \mapsto \underline{c}^k = X \underline{a}_k$$

Proposition \underline{c}^k est un vecteur propre de la matrice $XMX'D$ associé à la $k^{\text{ième}}$ plus grande valeur propre λ_k ; on a $\|\underline{c}^k\| = \sqrt{\lambda_k}$.

Démonstration

- $\|\underline{c}^k\|$ est l'écart-type du $k^{\text{ième}}$ facteur, égal à $\sqrt{\lambda_k}$.
- Montrons que les matrices $MX'DX$ et $XMX'D$ ont même ensemble de valeurs propres non nulles. Supposons $\lambda_k \neq 0$:

$$\begin{aligned} \text{a) } & MX'DX \underline{a}_k = \lambda_k \underline{a}_k \implies XMX'D(X \underline{a}_k) = \lambda_k (X \underline{a}_k) \\ \iff & XMX'D \underline{c}^k = \lambda_k \underline{c}^k \\ \text{b) } & XMX'D \underline{c}^k = \lambda_k \underline{c}^k \implies MX'DX(MX'D \underline{c}^k) = \lambda_k (MX'D \underline{c}^k). \end{aligned}$$

Toute valeur propre non nulle de l'une est valeur propre de l'autre; ces deux matrices ont donc même ensemble de valeurs propres non nulles. En outre, \underline{c}^k est vecteur propre de $XX'D$ associé à la valeur propre λ_k . ■

Proposition $\underline{v}^k = \frac{1}{\sqrt{\lambda_k}} \underline{c}^k = \frac{1}{\sqrt{\lambda_k}} X \underline{a}_k$ est solution du problème de la recherche de $\underline{v} \in \mathbb{R}^n$ tel que $\underline{v}'DXMX'D\underline{v}$ soit maximal sous les contraintes $\underline{v}'D\underline{v}^j = 0, \quad j = 1, \dots, k-1$ et $\underline{v}'D\underline{v} = 1$.

Démonstration

Notons $A = DXMX'D$. On cherche \underline{v} tel que $\underline{v}'A\underline{v}$ soit maximal sous les contraintes $\underline{v}'D\underline{v}^j = 0, \quad j = 1, \dots, k-1$ et $\underline{v}'D\underline{v} = 1$.

Toute solution de ce problème est vecteur propre de

$$D^{-1}A = D^{-1}(DXMX'D) = XMX'D$$

associé à la $k^{\text{ième}}$ plus grande valeur propre λ_k , comme \underline{c}^k . Mais, $\|\underline{c}^k\|^2 = \lambda_k$ (variance du $k^{\text{ième}}$ facteur). Donc, $\frac{\underline{c}^k}{\sqrt{\lambda_k}}$ est solution de ce problème. ■

Proposition Lorsque M est diagonale, d'éléments diagonaux M_{jj} , $\frac{1}{\sqrt{\lambda_k}} \underline{c}^k = \underline{v}^k$ est un vecteur unitaire du $k^{\text{ième}}$ axe principal de l'ACP non centrée des points $B^j(\underline{x}^j)$ affectés des poids M_{jj} dans (\mathbb{R}^n, D) .

Démonstration

On a établi que $\underline{v}^k = \frac{1}{\sqrt{\lambda_k}} \underline{c}^k$ était solution du problème de la recherche de $\underline{v} \in \mathbb{R}^n$ tel que $\underline{v}'DXMX'D\underline{v}$ soit maximal sous les contraintes $\underline{v}'D\underline{v}^j = 0, \quad j = 1, \dots, k-1$ et $\underline{v}'D\underline{v} = 1$.

Or, dans l'ACP du tableau X dans \mathbb{R}^p muni de la métrique M , la matrice des poids étant D , notée ACP (X, M, D) , on recherche pour $k = 1, \dots, r$, $\underline{u}_k \in \mathbb{R}^p$ qui rende maximal $\underline{u}'MX'DXM\underline{u}$ sous les contraintes $\underline{u}'M\underline{u}_j = 0, \quad j = 1, \dots, k-1$ et $\underline{u}'M\underline{u} = 1$.

Donc, on peut interpréter le premier problème comme étant l'ACP du tableau X' dans \mathbb{R}^n muni de la métrique D , la matrice des poids étant M , notée ACP (X', D, M) ; \underline{v}^k est alors un vecteur unitaire du $k^{\text{ième}}$ axe principal de cette ACP.

$$X' = \begin{pmatrix} \underline{x}^{1'} \\ \vdots \\ \underline{x}^{p'} \end{pmatrix} \begin{matrix} B^1 \\ \vdots \\ B^p \end{matrix}$$

L'origine de \mathbb{R}^n n'est pas le barycentre des points (B^j, M_{jj}) . On dit qu'on fait une ACP non centrée. ■

Pour l'ACP (X', D, M) , on définit les applications :

$$\begin{array}{l}
D : \quad \mathbb{R}^n \rightarrow \mathbb{R}^{n*} \\
\quad \underline{v}^k \mapsto D\underline{v}^k \\
\\
X' : \quad \mathbb{R}^{n*} \rightarrow \mathbb{R}^p \\
\quad D\underline{v}^k \mapsto X'D\underline{v}^k
\end{array}$$

4.3 Le schéma de dualité

En rassemblant les définitions des applications, on construit le schéma de dualité de l'ACP (X, M, D) :

$$\begin{array}{ccc}
(\underline{u}_k) \quad \mathbb{R}^p & \xleftarrow{X'} & \mathbb{R}^{n*} \\
\downarrow M & & \uparrow D \\
(\underline{a}_k) \quad \mathbb{R}^{p*} & \xrightarrow{X} & \mathbb{R}^n \quad (\underline{c}^k)
\end{array}$$

On lit sur ce schéma les données et les résultats principaux de l'ACP :

\underline{u}_k est vecteur propre de $X'DXM$;

$\underline{a}_k = M\underline{u}_k$ est vecteur propre de $MX'DX$;

$\underline{c}^k = XM\underline{u}_k$ est vecteur propre de $MX'D$.

On peut retrouver facilement les critères de détermination de $\underline{u}_k, \underline{a}_k, \underline{c}^k$.

Remarque $X'D\underline{v}^k = \frac{1}{\sqrt{\lambda_k}} X'D\underline{c}^k = \frac{1}{\sqrt{\lambda_k}} X'DXM\underline{u}_k = \frac{1}{\sqrt{\lambda_k}} \lambda_k \underline{u}_k = \sqrt{\lambda_k} \underline{u}_k$. ■

4.4 Analyse factorielle d'un tableau de distances

\underline{c}^k est vecteur propre de $\underbrace{XMX'}_{=W} D$.

$$XMX' = \begin{pmatrix} \underline{x}'_1 \\ \vdots \\ \underline{x}'_n \end{pmatrix} M (\underline{x}_1 \cdots \underline{x}_n)$$

L'élément (i, j) de W est $w_{ij} = \underline{x}'_i M \underline{x}_j = \langle \underline{x}_i, \underline{x}_j \rangle$ dans (\mathbb{R}^p, M) . Notons d_{ij} la distance entre les points A_i et A_j représentatifs des individus i et j dans \mathbb{R}^p :

$$\begin{aligned}
d_{ij}^2 &= \|\underline{x}_i - \underline{x}_j\|^2 = \|\underline{x}_i\|^2 - 2 \langle \underline{x}_i, \underline{x}_j \rangle + \|\underline{x}_j\|^2 \\
&= w_{ii} - 2w_{ij} + w_{jj}.
\end{aligned}$$

Notons $\overline{d}_i^2 = \sum_{l=1}^n p_l d_{il}^2$, $\overline{d}^2 = \sum_{l=1}^n \sum_{m=1}^n p_l p_m d_{lm}^2$.

Proposition On a $w_{ij} = \frac{1}{2}(\overline{d_i^2} + \overline{d_j^2} - d_{ij}^2 - \overline{d^2})$.

Démonstration

Soit \mathcal{J} l'inertie de l'ensemble des points (A_i, p_i) par rapport au barycentre.

$$\begin{aligned} \mathcal{J} &= \sum_{i=1}^n p_i \|\underline{x}_i\|^2 = \sum_{i=1}^n p_i w_{ii}. \\ \overline{d_i^2} &= \sum_{j=1}^n p_j d_{ij}^2 = w_{ii} + \mathcal{J}, \text{ car } \sum_1^n p_j \underline{x}_j = \underline{0}; \\ \overline{d_j^2} &= w_{jj} + \mathcal{J} \quad ; \quad \overline{d^2} = \sum_{i=1}^n p_i \overline{d_i^2} = 2\mathcal{J}; \\ \overline{d_i^2} + \overline{d_j^2} - d_{ij}^2 - \overline{d^2} &= 2w_{ij}. \blacksquare \end{aligned}$$

Donc, pour calculer la matrice $W = (w_{ij})$, il suffit de connaître les distances entre les points A_i .

Par conséquent, si on donne le tableau des distances (euclidiennes) entre les A_i , on peut construire la matrice W , donc WD , et on peut calculer les composantes principales. On peut alors faire les représentations graphiques des projections des points A_i sur les axes principaux.

4.5 Le cercle des corrélations

Notons y^j le caractère réduit associé au caractère $x^j : y^j = \frac{x^j}{s_j}$. Dans $\mathbb{R}^n : \underline{y}^j = \frac{\underline{x}^j}{\|\underline{x}^j\|} ; \|\underline{y}^j\| = 1$.

1. Notons $C^j(y^j)$ le point de \mathbb{R}^n représentatif du caractère y^j .

1) Dans \mathbb{R}^n , les points C^j sont situés sur la sphère de centre l'origine et de rayon 1. $\cos(\underline{x}^l, \underline{x}^m) = \cos(\underline{y}^l, \underline{y}^m)$ représente le coefficient de corrélation linéaire entre les caractères x^l et x^m .

Si ces caractères sont fortement corrélés positivement (cosinus proche de 1 \iff angle proche de 0), alors les points C^l et C^m sont proches.

Si ces caractères sont fortement corrélés négativement (cosinus proche de -1 \iff angle proche de π), alors les points C^l et C^m sont presque diamétralement opposés.

2) Soit un plan passant par l'origine et engendré par les vecteurs \vec{v}^l et \vec{v}^m (dans le cas où M est diagonale, c'est un plan principal de l'ACP des points B^j).

L'intersection de la sphère de rayon 1 et de ce plan est un cercle de rayon 1 appelé *cercle des corrélations*.

Si, sur ce plan, les projections de C^l et de C^m sont proches du cercle, ces points sont bien représentés par projection.

Si, en outre, ces projections sont proches l'une de l'autre (respectivement presque diamétralement opposées), ceci traduit que les points C^l et C^m sont proches (respectivement presque diamétralement opposés), donc que les caractères x^l et x^m sont fortement corrélés positivement (respectivement négativement).

3) La projection de C^j sur l'axe $(0, \vec{v}^k)$ (on note O l'origine de \mathbb{R}^n) a pour abscisse $\langle \underline{y}^j, \underline{v}^k \rangle$: c'est la covariance de y^j et de v^k , donc le coefficient de corrélation linéaire du caractère x^j et du $k^{\text{ième}}$ facteur.

Donc, dans le plan de projection $(0, \vec{v}^1, \vec{v}^j)$, on représente le caractère x^j par un point qui a pour coordonnées les coefficients de corrélation linéaire de ce caractère avec les deux premiers facteurs. On fait de même pour tout autre plan de projection.

On calcule ces coefficients de corrélation linéaire :

$$\langle \underline{y}^j, \underline{v}^k \rangle = (\underline{y}^j)' D \underline{v}^k = \frac{1}{s^j} \frac{1}{\sqrt{\lambda_k}} (\underline{x}^j)' D \underline{c}^k = \frac{1}{s^j} \frac{1}{\sqrt{\lambda_k}} (\underline{x}^j)' D X M \underline{u}_k.$$

$$\text{Or: } X' = \begin{pmatrix} \underline{x}^{1'} \\ \vdots \\ \underline{x}^{j'} \\ \vdots \\ \underline{x}^{p'} \end{pmatrix} ; (\underline{x}^j)' D X M \underline{u}_k \text{ est la } j^{\text{ième}} \text{ ligne de } X' D X M \underline{u}_k = \lambda_k \underline{u}_k.$$

$$\text{Donc : } \langle \underline{y}^j, \underline{v}^k \rangle = \sqrt{\lambda_k} \frac{u_k^j}{s^j}.$$

Proposition *Le coefficient de corrélation linéaire du caractère x^j et du $k^{\text{ième}}$ facteur est égal à $\sqrt{\lambda_k} \frac{u_k^j}{s^j}$.*

Remarque La mesure algébrique de la projection de \vec{e}_j sur le $k^{\text{ième}}$ axe principal (G, \vec{u}_k) (dans \mathbb{R}^p) est égale, dans le cas où M est diagonale, à $M_{jj} u_k^j$.

Le rapport du coefficient de corrélation linéaire du caractère x^j et du $k^{\text{ième}}$ facteur et de cette mesure algébrique est égal à $\frac{\sqrt{\lambda_k}}{M_{jj} s^j}$.

Dans le cas de l'ACP normée, lorsque l'on prend des données centrées réduites et $M = I$, ce rapport est égal à $\sqrt{\lambda_k}$ et ne dépend pas de j . C'est pourquoi on superpose parfois dans ce cas la projection des points C^j et la projection des points A_i . ■

5 Interprétation des résultats d'une ACP

On représente la projection des points-individus A_i et des axes-caractères (G, \vec{e}_j) sur le plan principal $(G, \vec{u}_1, \vec{u}_2)$ et éventuellement sur d'autres plans principaux. On représente les cercles de corrélation correspondants.

On peut diviser le plan d'interprétation en quatre parties.

5.1 Axes principaux à conserver

On note la qualité globale de représentation sur chaque axe (appelée aussi pourcentage d'inertie expliquée par l'axe ou pourcentage de variance expliquée par le facteur correspondant). Pour décider quels axes conserver, on peut s'appuyer sur les trois principes suivants, qui peuvent être contradictoires :

1. On conserve tout axe que l'on peut interpréter.
2. On arrête l'introduction des axes lorsque la variance d'un facteur correspondant à un axe est inférieure à celle d'un caractère.

Par exemple, en *ACP* normée, lorsqu'on utilise des caractères réduits, on arrête l'introduction lorsque la variance est plus petite que 1.

3. On arrête l'introduction lorsque la qualité de représentation sur un axe est beaucoup plus faible que celle du précédent.

Exemple Axe 1 : 30 % - Axe 2 : 20 % - Axe 3 : 18 % ; on peut conserver l'axe 3.

Axe 1 : 30 % - Axe 2 : 20 % - Axe 3 : 6 % ; on peut ne pas conserver l'axe 3.

5.2 Etude des corrélations entre les caractères initiaux

On étudie le cercle des corrélations et la matrice des corrélations. On détermine les caractères fortement corrélés positivement et ceux fortement corrélés négativement.

- Lorsque deux caractères sont représentés par des *points proches du cercle et proches l'un de l'autre*, il existe une forte corrélation linéaire positive entre eux ;
- lorsque deux caractères sont représentés par des *points proches du cercle et presque diamétralement opposés*, il existe une forte corrélation linéaire négative entre eux.

On lit les coefficients dans la matrice des corrélations.

Remarque Des caractères fortement corrélés positivement varient dans le même sens et peuvent représenter le même phénomène. Si l'on conserve trop de caractères fortement corrélés positivement, la différence entre deux individus, qui seraient semblables vis-à-vis de ces caractères, mais dissemblables vis-à-vis des autres, peut apparaître fortement atténuée. Pour remédier à ceci, on peut éliminer certains de ces caractères ou leur donner un poids qui diminue leur importance dans l'étude, ce qui revient à changer de métrique M . ■

5.3 Interprétation des facteurs

Que représentent les combinaisons linéaires des caractères initiaux centrés que l'on a déterminées ? On peut faire cette interprétation en s'appuyant sur l'une ou l'autre des deux analyses suivantes.

5.3.1 Analyse des corrélations des facteurs avec les caractères

On détermine les caractères fortement corrélés positivement d'une part, négativement d'autre part, avec les facteurs. Ceci peut permettre de donner une signification à un facteur, en ayant déterminé les caractères qui varient en moyenne dans le même sens que ce facteur et ceux qui varient en moyenne en sens contraire.

Exemple Dans une analyse où l'on a mesuré des variables économiques sur les pays du monde, le premier facteur peut être corrélé positivement avec des caractères tels que le PNB par habitant, les importations, les exportations, le pourcentage d'actifs dans le secteur tertiaire et corrélé négativement avec des variables concernant l'importance relative du secteur primaire (agriculture, sylviculture, pêche) ; l'axe principal correspondant oppose alors des pays très développés économiquement (qui ont des valeurs fortes pour les caractères du premier groupe et des valeurs faibles pour les caractères du deuxième) à des pays moins développés (qui sont dans le cas contraire) : on peut dire que le premier facteur représente le développement économique.■

Remarque *Cas d'un facteur de taille* : c'est un facteur tel que tous les caractères sont fortement corrélés positivement (ou tous négativement) avec lui. L'axe principal correspondant oppose les individus ayant de fortes valeurs pour tous les caractères à ceux ayant de faibles valeurs. On peut parfois remédier au phénomène de taille en *transformant les caractères*, ici en remplaçant des caractères de taille par des caractères de structure. par exemple :

1) On fait une étude sur les départements français. Le caractère "nombre de médecins" dépend de la population du département : on peut l'interpréter comme un caractère de taille. On peut alors le remplacer par le caractère "nombre de médecins pour 1 000 habitants" (ou "densité médicale"), qui est un caractère de structure.

2) On fait une étude sur les notes d'étudiants dans différentes matières. Il peut apparaître que certains ont de bonnes notes partout, alors que d'autres en ont de mauvaises partout ; le premier facteur peut alors être un facteur de taille, l'axe principal opposant les étudiants ayant de bonnes notes à ceux ayant de mauvaises notes.■

5.3.2 Analyse de la position d'individus ayant une contribution importante

Si certains individus ont une contribution très importante à l'inertie expliquée par un axe principal et un type bien défini, l'étude de leur position par rapport à cet axe peut permettre de donner une signification au facteur correspondant. Certains individus peuvent parfois expliquer à eux seuls la construction d'un axe principal.

Remarque Certains individus ayant une contribution très importante peuvent être considérés, après examen des valeurs des caractères, comme aberrants. On peut alors les éliminer de la détermination des axes principaux, et cependant faire figurer leur projection sur les plans (*individus supplémentaires ou non actifs*).■

Exemple On fait une étude sur les départements français. Il peut apparaître que les départements d'outre-mer ont des valeurs des caractères très différentes de celles des autres. Après avoir étudié leurs caractéristiques, on peut dans ce cas limiter éventuellement l'analyse aux départements de la métropole.■

Une analyse de données se fera donc généralement en plusieurs étapes : après un premier passage, on interprète les résultats et on élimine éventuellement des caractères ou des individus ; on effectue alors un deuxième passage et on reprend le même procédé.

5.4 Interprétation de la position des projections des points-individus

On interprète la position par rapport aux axes principaux de la projection des points représentatifs des individus.

Si un individu a une abscisse positive (respectivement négative) élevée en valeur absolue sur un axe principal, c'est qu'il a une forte (respectivement faible) valeur du facteur correspondant et, en général, une forte (respectivement faible) valeur des caractères fortement corrélés positivement avec ce facteur et une faible (respectivement forte) valeur des caractères fortement corrélés négativement avec ce facteur. On confirme ceci en consultant le tableau des données.

La proximité entre deux individus sur un plan principal ne traduit pas nécessairement une ressemblance entre ces individus vis-à-vis de tous les caractères, en particulier si la qualité de représentation de ces individus sur le plan considéré est faible, mais traduit une ressemblance vis-à-vis des facteurs correspondants.

Si le nombre d'individus est trop important, ou si les individus constituent une population ou un échantillon représentatif de cette population pour laquelle l'étude des résultats individuels ne présente pas d'intérêt, alors on peut regrouper les individus en *classes*

- empiriquement, à partir des graphiques de l'ACP ;
- automatiquement, par une méthode de *classification*.

On interprète alors la position des classes par rapport aux axes principaux.

Exemple Dans une banque, on peut étudier la population des clients et vouloir segmenter cet ensemble en classes dont on étudie les caractéristiques globales sans étudier celles de chaque client.■

Remarque Dans le logiciel SPAD, la classification est faite non pas à partir des valeurs des caractères, mais à partir des valeurs des facteurs. L'utilisateur déclare le nombre de facteurs qu'il souhaite introduire. Après la classification, on peut représenter automatiquement les classes sur les plans d'axes principaux, le code d'un individu étant remplacé sur la figure par le numéro de la classe.■

Remarque *Représentation d'un caractère qualitatif* : on peut représenter un caractère qualitatif sur un plan principal. On représente chaque modalité du caractère par un point

qui est le barycentre des projections sur ce plan des points représentatifs des individus ayant cette modalité. Si ce caractère est ordinal, on peut joindre ces points par des flèches orientées en respectant l'ordre des modalités.■

Réunion du 10 février 2006 : Docteur Elisabeth Luporsi

Après un rappel sur les essais cliniques, le problème exposé était le suivant : Actuellement il faut un grand nombre de sujets pour les essais cliniques, en effet, le bénéfice absolu des molécules est faible avec un intervalle de confiance large, ce qui veut dire que certains patients bénéficient du traitement et d'autres pas. On peut se demander si le résultat trouvé d'amélioration de la survie sans récurrence est lié au médicament testé.

L'objectif est de sélectionner les patients qui doivent répondre à un traitement et nous souhaitons enlever la survie globale dans le cadre des essais de Phase III, ne plus parler que de survie sans récurrence et de réaliser des essais de phase II avec marqueurs intermédiaires de réponse ; en phase II l'objectif est la toxicité (la dose est définie par les essais de phase I qui précisent la DLT (dose limitante tolérable). Ces essais de Phase II (avec marqueurs intermédiaires de réponse) pourraient utiliser des méthodes bayésiennes (couplant dans un même objectif la réponse et la tolérance) et nécessiteraient un nombre plus important de patients (150 à 200 patients) que actuellement (20 à 30 patients). Le problème actuel est l'inflation du risque alpha (d'ordinaire de 5%) du fait d'objectifs secondaires nombreux et pour lesquels le nombre de sujets n'a pas été calculé (car calcul toujours fait sur l'objectif principal)

Par ailleurs a été développé la problématique de la protéomique et de la génomique qui actuellement ne permet pas de sélectionner les gènes car il existe 95% de faux positifs.

Parameter Estimation of Pharmacokinetics Models in the presence of Timing Noise

Thierry Bastogne*, Sophie Mézières-Wantz†, Nacim Ramdani‡, Pierre Vallois§
and Muriel Barberi-Heyob¶

February 15, 2007

Abstract

The problem addressed in this paper deals with the parameter estimation of *in vitro* uptake kinetics of drugs into living cells in presence of timing noise. Effects of the timing noise on the bias and variance of the output error are explicitly determined. A bounded-error parameter estimation approach is proposed as a suited solution to handle this problem. Application results are presented and emphasize its effectiveness in such an experimental framework.

Keywords parameter estimation, pharmacokinetics models, timing errors, bounded errors, biological systems.

*T. Bastogne is with the Centre de Recherche en Automatique de Nancy (CRAN), Nancy-Université, CNRS, BP 239, F-54506 Vandœuvre-lès-Nancy Cedex, France thierry.bastogne@cran.uhp-nancy.fr

†S. Mézières-Wantz is with the Institut de Mathématiques Elie Cartan, Nancy-Université, CNRS, BP 239, F-54506 Vandœuvre-lès-Nancy Cedex, France sophie.wantz@iecn.u-nancy.fr

‡N. Ramdani was with the CERTES EA 3481 Université Paris 12 Val de Marne. He is now with the Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier (LIRMM) UMR CNRS 5506 Université de Montpellier 2, 161 rue Ada, 34392 Montpellier Cedex 5 - France. nacim.ramdani@lirmm.fr

§P. Vallois is with the Institut de Mathématiques Elie Cartan, Nancy-Université, CNRS, BP 239, F-54506 Vandœuvre-lès-Nancy Cedex, France pierre.vallois@iecn.u-nancy.fr

¶M. Barberi-Heyob is with the Centre de Recherche en Automatique de Nancy (CRAN), Nancy-Université, CNRS, Centre Alexis Vautrin, 54511 Vandœuvre-lès-Nancy Cedex, France m.barberi@nancy.fnclcc.fr

1 Introduction

Pharmacokinetics is concerned with both the absorption speed of a drug and how the body processes a drug. In chemical kinetics, reactions are generally described by differential equations which link the reaction rate with concentrations or pressures of reactants. In molecular cell biology, because of the complexity of systems, the nature of some reactions is still unclear. This paper is focused on the intracellular uptake kinetics of a photosensitizing drug (PS), *i.e.* the rate of how photosensitizing molecules are absorbed by living cancer cells [1]. The delivery control of the photosensitizing agent into the cancer cells is one the major factor on the therapeutic efficiency of the photodynamic therapy (PDT) [2–5]. Most of the PS uptake kinetics models are non-parametric, the temporal evolution of the PS intracellular concentration is described by step responses. The purpose of this study is the estimation of kinetics model parameters from data collected during kinetics experiments. Few papers have been reported for the application of system identification techniques to pharmacokinetics modelling problems [6–8].

Unfortunately, these approaches cannot be applied to the PS uptake kinetics. The intracellular PS concentration ($[P_i]$) is measured by a spectrofluorimeter but the PS reacts with the light sent by the measurement system. As a consequence, $[P_i]$ can only be measured at one time instant t_j . In other terms, a biological sample cannot be used for several measurements. Collecting N_t data points of the kinetics then requires to repeat N_t times the same experiment (N_t biological samples) with identical initial conditions. To avoid a too high cost of such an experiment set up, N_t is generally kept

small, *i.e.* $N_t \leq 10$. This limitation on N_t is the first problem to overcome for estimating kinetics parameters. The second difficulty is the low signal-to-noise ratio. The latter is due to a great measurement variability when working on living cells which are very sensitive to external disturbances. Thirdly, the choice of the stimulus signal is restricted to step signals which correspond to the amount of PS injected into the culture medium wells at time $t = 0$. A last issue is a timing offset error in the measurement samples. This timing error is bounded and can reach until $\pm 15mn$ for a measurement time sequence $\{t_j\} = \{1, 2, 4, 6, 8, 14, 18, 24h\}$. This error cannot be neglected and has rarely been taken into account in parameter estimation problems [9], particularly in this restricted experimental framework.

The problem addressed in this paper deals with the estimation of kinetics model parameters in presence of timing noise. Parameter estimation algorithms in system identification methods are often based on the minimization of a quadratic function of the output error, *i.e.* the difference between the system and the model outputs. The sensitivity of the output error to timing noise comparatively to input and output noise is unknown. The contributions of this paper are twofold:

- the asymptotic effects of the timing noise on the output error are explicitly determined. They are compared with the ones of input and output noise. These results are obtained by assuming that there is no modeling error between the model and the biological system;
- a solution suited to this system identification problem is proposed and applied to *in vitro* data in a second part of this paper in order to assess its effectiveness in practice. The identification problem is addressed in the *bounded-error* context and is solved with a set projection algorithm based on interval analysis introduced in [10].

2 Experimental set up

Fig. 1 presents the *in vitro* experimental set up associated with the uptake kinetics of a photosensitizing drug into living cells. Cells are inoculated in $250\mu l$ culture wells and are exposed at time $t = 0$

Table 1: Main notations

Symb.	Description
t	time variable
t_j	theoretical time instant of the j^{th} sample
\hat{t}_j	real time instant of the j^{th} sample
$u(t)$	noise-free input signal
$\hat{u}(t)$	real input signal
$y(t)$	system output variable
$\hat{y}(t)$	measured output variable
$y_M(t)$	model output variable
n_u	input noise
n_y	output noise
n_t	timing noise
$e_y(t)$	output error variable
\mathcal{S}	biological system
$\mathcal{M}(p)$	parametric model
p	vector of parameters
$N_t = \text{card}(\{t_j\})$	number of data points
N_r	number of repeated exp. at each time instant
x'	transposition of x
$\mathcal{N}(\mu, \sigma)$	gaussian distribution with μ : mean σ : standard deviation
$\mathcal{U}(a, b)$	uniform distribution on the interval $[a; b]$
$\mathbb{E}\{\cdot\}$	math. expectation operator

to a photosensitizing drug D with a given concentration $[D]$. The input variable $u(t)$ is a step signal which corresponds to the amount of drug injected into the well at time $t = 0$. The output variable $y(t)$ is the amount of drug absorbed by the cells. $\hat{y}(t)$ is the measurement of $y(t)$ given by a spectrofluorimeter at times $\{t_j\}$ avec $j \in \{1, \dots, N_t\}$. Because the spectrofluorimeter affects the biological state of the photosensitizing drug, each measurement requires at least one culture well. Moreover, N_r identical wells are handled by the experimenter at each time instant t_j to *a posteriori* estimate the repeatability of the measurements. Globally, $N_r \times N_t$ wells are handled during the whole experiment. All the wells are prepared in the same initial conditions.

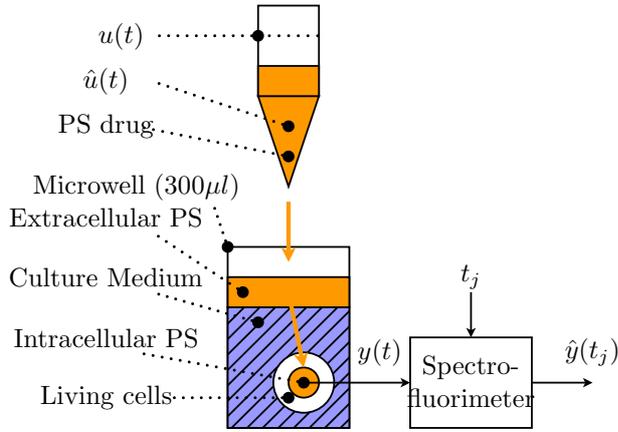


Figure 1: *in vitro* experimental set up

3 Model and noise descriptions

The determination of a parametric model describing the uptake kinetics of a photosensitizing drug into living cells by extracting information from observations of u and y is a system identification problem [11–13]. In the sequel, it is assumed that the system and the model are identical $\mathcal{M}(p) \equiv \mathcal{S}(p)$. However, as depicted in Fig. 2, three kinds of uncertainties are examined and are represented by output, input and timing noise. Output and input noise (n_y and n_u) are described by stationary stochastic processes added to the output and input signals. The timing noise (n_t) is a sequence of random variables added to the timing sequence $\{t_j\}$ controlling the sampling process of the output signal. \hat{t}_j is the real time instant at which the output variable y is measured while t_j represents the theoretical measurement time instant noted by the experimenter. $e_y(t_j)$ denotes the output error between the system and the model outputs ($\hat{y}(t_j)$ and $y_{\mathcal{M}}(t_j)$ respectively).

3.1 Model structure

In [14], it is shown that a first-order transfer function is a parsimonious model structure for describing the uptake kinetics of the *chlorin e6* photosensitizing drug into HT29-A4 cancer cells (human colon cancer cell line). For the sake of simplicity, it is assumed in the sequel that $\mathcal{M}(p)$ and $\mathcal{S}(p)$ both rely

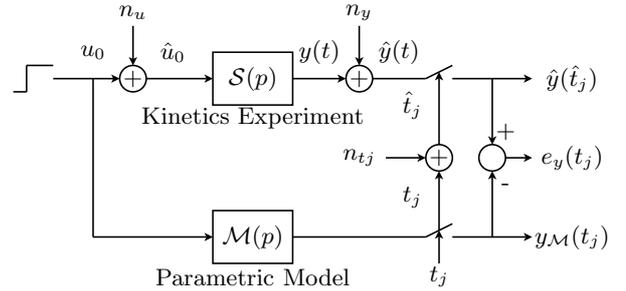


Figure 2: Description of uncertainties

on a first-order transfer function defined as follows

$$\mathcal{S}: \quad T \cdot \frac{dy}{dt} + y(t) = k \cdot \hat{u}(t) \quad (1)$$

$$\mathcal{M}: \quad T \cdot \frac{dy_{\mathcal{M}}}{dt} + y_{\mathcal{M}}(t) = k \cdot u(t), \quad (2)$$

with $y(0) = y_{\mathcal{M}}(0) = 0$. T and k denote the time constant and the static gain respectively, and $u(t)$ is a step signal of magnitude u_0 defined in (7). As a result, the intracellular concentration of the photosensitizing drug y follows a mono-exponential kinetics defined by

$$y(t) = k \cdot \hat{u}_0 \cdot (1 - e^{-t/T}). \quad (3)$$

3.2 Output noise

Conjugated effects of measurement noise and disturbances are usually described by a stochastic variable n_y added to y such that

$$\hat{y}(t_j) = y(t_j) + n_y(t_j), \quad (4)$$

where y and \hat{y} denote the real biological response and its measurement respectively. Herein, it is assumed that $\{n_y(t_j)\}$ is an independent and identically distributed sequence of gaussian variables with

$$n_y(t_j) = \sigma_y \cdot G_y^j, \quad (5)$$

where σ_y denotes the standard deviation of n_y and $G_y^j \approx \mathcal{N}(0, 1)$.

3.3 Input noise

u is a step signal defined by

$$u(t) = \begin{cases} 0 & t < 0 \\ u_0 & t \geq 0 \end{cases} \quad (6)$$

The step magnitude (u_0) represents the amount of the injected drug. The duration of injection is not significant compared to the duration of the experiment. The drug administration is usually carried out by multichannel micropipettes. For technical reasons, the real filling levels of drug in the cones are not identical and do not match with the dose stipulated by the experimenter. This error is represented by an input noise n_u added to u such that

$$\hat{u}(t) = \begin{cases} 0 & t < 0 \\ \hat{u}_0 = u_0 + n_u & t \geq 0 \end{cases} \quad (7)$$

where u_0 and \hat{u}_0 denote the *a priori* prescribed dose and the effectively administered dose respectively. The input noise is defined by

$$n_u = \sigma_u \cdot G_u, \quad (8)$$

where σ_u denotes the standard deviation of n_u and $G_u \approx \mathcal{N}(0, 1)$. In parameter estimation, output and input noise have been taken into account by error-in-variables approaches [15].

3.4 Timing noise

Each experimentation is carried out in a sterile framework in which each well is carefully handled by the experimenter. For each time instant t_j of the experimental set up, the living cells have to be washed, removed, lysed and diluted in ethanol by the experimenter. Such operations can take about 20 to 30mn during which each well, among the N_r ones, is individually handled. In practice, only the time instants t_j^- and t_j^+ corresponding to the beginning and the end of the experiment are noted by the experimenter. The nominal measurement time instant t_j noted by the experimenter in his table is an average time instant defined by $t_j = (t_j^+ + t_j^-)/2$. The real time instant \hat{t}_j at which the uptake kinetics is stopped and measured, is unknown. This lack of precision in the timing of experiments is described by a timing noise n_t . The latter is a bounded error whose bounds are measured. In a first part, the timing noise sequence $\{n_{tj}\}$ is supposed to independent and identically distributed such that

$$n_{tj} \approx \mathcal{U}(t_j^-, t_j^+), \quad (9)$$

or

$$n_{tj} = -\frac{\tau}{2} + \tau \cdot U_t^j, \quad (10)$$

with $\tau = t_j^+ - t_j^-$ and $U_t^j \approx \mathcal{U}(0, 1)$. τ denotes the width of the timing uncertainty interval.

4 Stochastic modeling

Given the previous assumptions about the input, output and timing noise, the expression of $e_y(t_j)$ becomes

$$\begin{aligned} e_y(t_j) &= \hat{y}(\hat{t}_j) - y_{\mathcal{M}}(t_j) & (11) \\ &= k \cdot (u_0 + \sigma_u G_u) \cdot (1 - e^{-\frac{1}{T}(t_j - \frac{\tau}{2} + \tau U_t^j)}) + \sigma_y G_y^j \\ &\quad - k \cdot u_0 \cdot (1 - e^{-\frac{t_j}{T}}), & (12) \end{aligned}$$

where k , T are given.

The mathematical expectation of $e_y(t_j)$ is defined in Proposition 4.1, its demonstration is developed in appendix A.

Proposition 4.1

$$\mathbb{E}\{e_y(t_j)\} = k \cdot u_0 \cdot e^{-\frac{t_j}{T}} \left(1 - \operatorname{sinhc}\left(\frac{\tau}{2T}\right)\right), \quad (13)$$

with $\operatorname{sinhc}(x) = \sinh(x)/x$ denotes the hyperbolic sinus cardinal function of x .

Since $x \rightarrow \operatorname{sinhc}(x)$ is increasing on \mathbb{R}^+ , equation (13) shows that $\mathbb{E}\{e_y(t_j)\} < 0$. This systematic bias is only due to the timing noise. The absolute value of the mean output error increases with respect to τ and is null only for $\tau = 0$.

Let us introduce a new random variable $\varepsilon_y^{\%}(t_j) = 100 \cdot e_y(t_j)/y_{\mathcal{M}}(t_j)$. It can be directly deduced from (13) that

$$\mathbb{E}\{\varepsilon_y^{\%}(t_j)\} = 100 \cdot \frac{1 - \operatorname{sinhc}\left(\frac{\tau}{2T}\right)}{e^{\frac{t_j}{T}} - 1} \quad (14)$$

$\mathbb{E}\{\varepsilon_y^{\%}(t_j)\}$ represents the mean relative output error. Fig. 3 confirms the previous remarks concerning the evolution of the mean relative output error with respect to τ . Moreover, it is shown that for $\tau = 0.5h$ (*i.e.* 30mn, a realistic value in practice) $\mathbb{E}\{\varepsilon_y^{\%}(t_j)\}$ reaches about 30% of the output variable for $t_j = 1h$ and about 12% of the output variable for $t_j = 2h$. For $t_j = 8h$, the effect of $\mathbb{E}\{\varepsilon_y^{\%}(t_j)\}$ becomes negligible.

The variance of $e_y(t_j)$ is defined in Proposition 4.2, its demonstration is developed in appendix B.

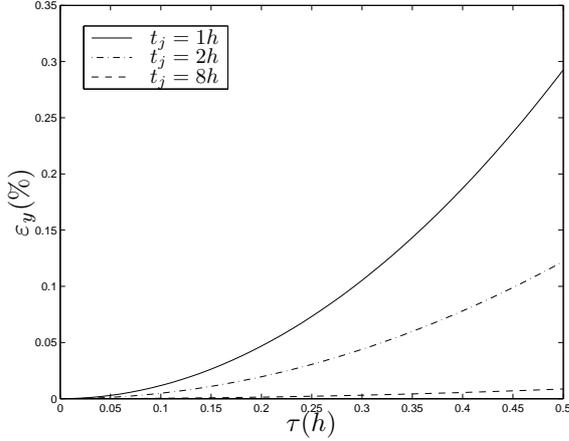


Figure 3: Mean relative output error

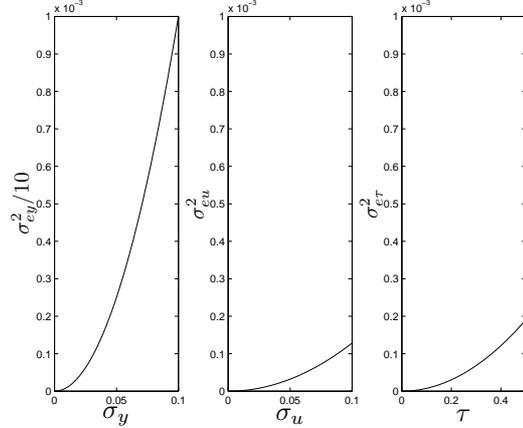


Figure 4: Variance analysis of the output error

Proposition 4.2

$$\text{Var}\{e_y(t_j)\} = \sigma_y^2 + k^2\sigma_u^2 + k^2e^{-\frac{t_j}{T}} \cdot \text{sinhc}\left(\frac{\tau}{2T}\right)(A + B) \quad (15)$$

$$\text{with: } A = e^{-\frac{t_j}{T}} \left(\cosh\left(\frac{\tau}{2T}\right) - \text{sinhc}\left(\frac{\tau}{2T}\right) \right) u_0^2 \text{ and } B = \left(e^{-\frac{t_j}{T}} \cosh\left(\frac{\tau}{2T}\right) - 2 \right) \sigma_u^2.$$

Let us compare the values taken by $\text{Var}\{e_y(t_j)\}$ for each kind of noise. Contributions of output, input and timing noise are described by three variances: σ_{ey}^2 , σ_{eu}^2 and $\sigma_{e\tau}^2$ defined as follows

- $\sigma_y \neq 0, \sigma_u = 0, \tau = 0$:

$$\text{Var}\{e_y(t_j)\} = \sigma_y^2 = \sigma_{ey}^2, \quad (16)$$

- $\sigma_y = 0, \sigma_u \neq 0, \tau = 0$:

$$\text{Var}\{e_y(t_j)\} = k^2\sigma_u^2 \left(e^{-\frac{t_j}{T}} - 1 \right)^2 = \sigma_{eu}^2, \quad (17)$$

- $\sigma_y = 0, \sigma_u = 0, \tau \neq 0$:

$$\begin{aligned} \text{Var}\{e_y(t_j)\} &= k^2u_0^2e^{-\frac{2t_j}{T}} \left(\cosh\left(\frac{\tau}{2T}\right) \right. \\ &\quad \left. - \text{sinhc}\left(\frac{\tau}{2T}\right) \right) \text{sinhc}\left(\frac{\tau}{2T}\right) \\ &= \sigma_{e\tau}^2. \end{aligned} \quad (18)$$

Fig. 4 presents the evolution $\sigma_{ey}^2/10$, σ_{eu}^2 and $\sigma_{e\tau}^2$ with respect to σ_y , σ_u and τ respectively. Parameters of simulations are $k = 1$, $T = 3$, $\text{SNR}_u^{\text{min}} \approx$

20dB , $\text{SNR}_y^{\text{min}} \approx 20\text{dB}$ and $0 \leq \tau \leq 30\text{mn}$. These values have been chosen from empirical knowledge of biologists and experimental results reported in [16]. Fig. 4 points out that the output noise is more significant in the sense of the output error variance than the input and timing noise in these simulation conditions.

5 Bounded-error estimation with *in vitro* data

Bounded-error or set-membership approaches allow to estimate parameters and their uncertainty in inverse problem contexts in which all uncertain quantities are assumed as unknown but bounded with known bounds. No further hypothesis about probability distributions is stated. Several algorithms have been developed to solve estimation problems stated in the bounded-error context. When models are non linear, most approaches use interval analysis and constraint propagation techniques. One important feature of these techniques is the fact that they derive guaranteed computations, i.e. they derive a numerical proof of property or non-property. They are rather mature techniques and have already been successfully applied for solving problems in biology, chemical or thermal engineering, economics, computer vision or robotics, when guaranteed computations were essential [10, 17].

In this part, we assume that all the uncertain quantities satisfy this *bounded-error* property and

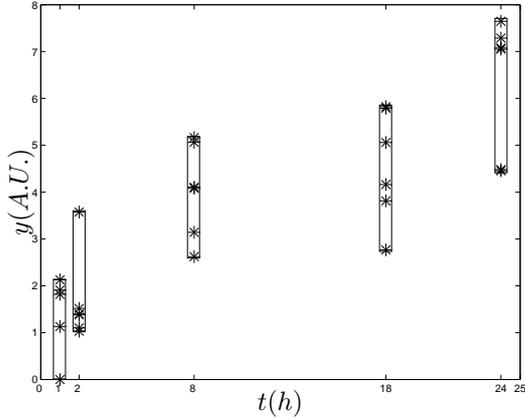


Figure 5: Experimental data, intervals and boxes

bounded-error estimation techniques are applied to experimental data collected during *in vitro* uptake kinetics experiments of a PS into human malignant glioma cells. The experimental protocol is defined by $N_t = 5$, $N_r = 6$ and $u_0 = 25mol$.

Table 2: Prior feasible intervals for the data

j	t_j (h)	$[t_j]$	$[\hat{y}(t_j)]$
1	1	[0.67;1.33]	[0;0.607]
2	2	[1.67;2.33]	[0.238;0.861]
3	8	[7.67;8.33]	[0.681;1.396]
4	18	[17.67;18.33]	[0.661;1.447]
5	24	[23.67;24.33]	[1.376;2.459]

Fig. 5 presents the experimental data of six PS uptake kinetics carried out in the same experimental framework. Each cross corresponds to one measurement. The output variable measured by the spectrofluorimeter is given in arbitrary unit. Prior intervals $[\hat{y}(t_j)]$ and $[t_j]$ on the output measurements and the time instants are given in table 2. Bounds of $[t_j]$ have been measured during the kinetics experiment. $[\hat{y}(t_j)]$ has been determined from the minimum and maximum values of measurements. The uncertainty associated with each pair of output and time data is materialized by a box. The set $\hat{\mathbb{P}}$ to be characterized consists of all the values of $\mathbf{p} = (k, T)'$ such that the graph of the function

$$g(\mathbf{p}, t) = k \cdot u_0 \cdot (1 - e^{-t/T}) \quad (19)$$

goes through all five boxes of Fig. 5. $\hat{\mathbb{P}}$ is defined

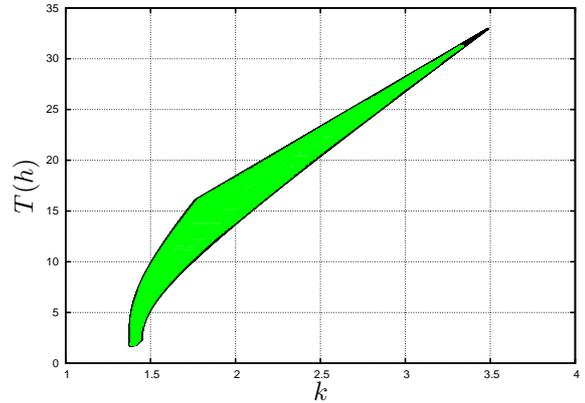


Figure 6: *a posteriori* estimate of the parameter set $\hat{\mathbb{P}}$

as

$$\hat{\mathbb{P}} = \{\mathbf{p} \in \mathbb{R}^{+2} | \exists \mathbf{t} \in \mathbb{R}^{+5}, (g(\mathbf{p}, \mathbf{t}), \mathbf{p}, \mathbf{t}) \in \mathbb{X}\} \quad (20)$$

with

$$g(\mathbf{p}, \mathbf{t}) = (g(\mathbf{p}, t_1) \quad \dots \quad g(\mathbf{p}, t_{N_t})), \quad (21)$$

and

$$\mathbb{X} = [\hat{y}(t_1)] \times \dots \times [\hat{y}(t_5)] \times [\check{k}] \times [\check{T}] \times [t_1] \times \dots \times [t_5], \quad (22)$$

and the prior box for the parameters given by

$$\check{\mathbb{P}} = [\check{k}] \times [\check{T}] = [1, 4] \times [1, 40]. \quad (23)$$

$\hat{\mathbb{P}}$ can be estimated in a guaranteed way using a set inversion algorithms based on parameter space partitioning, interval analysis and constraint propagation techniques (see [10] and the references therein). Fig. 6 presents the estimate of $\hat{\mathbb{P}}$ when the partitioning algorithm is set not to partition boxes with a size smaller than 0.01

In Fig. 6, the paving form associated with $\hat{\mathbb{P}}$ is composed of grey and black boxes. The grey boxes have been proved to be included in \mathbb{P} but no conclusion has been reached for the black ones. The external envelope of $\hat{\mathbb{P}}$ is defined by $\hat{k} \in [1.37; 3.49]$ and $\hat{T} \in [1.7; 33]$. This results shows that the estimation uncertainty on the time-constant is larger than the one on the static gain. Fig. 7 depicts the *a posteriori* estimate of the output set $\hat{\mathbb{Y}}$, a set of time trajectories defined by

$$\hat{\mathbb{Y}} = \{(t, y) \in \mathbb{R}^+ \times \mathbb{R} | y(t) = k u_0 (1 - e^{-t/T}), \text{ with } (k, T)^T \in \hat{\mathbb{P}}\}. \quad (24)$$

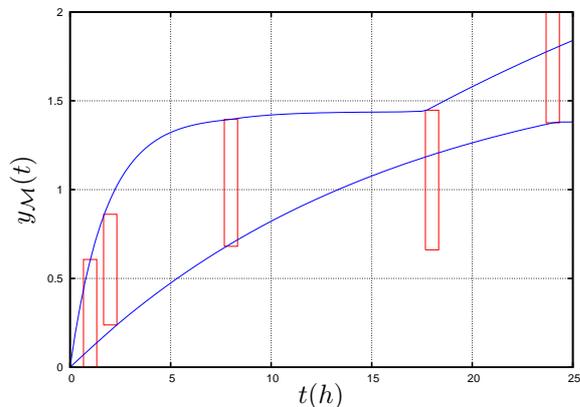


Figure 7: *a posteriori* estimate of the output set \hat{Y}

This figure points out a wide variation of the initial slope of the step response which explains the large uncertainty on the time-constant estimate. In this study case, the wide variation of the initial slope is mainly due to the height of the boxes rather than their width. In other terms, in this application, the uncertainty on the time-constant estimate is mainly caused by the output noise rather than the timing noise.

6 Conclusion

This paper firstly investigates the asymptotic consequences of timing errors in collected data on the parameter estimation of kinetics models and more precisely their effects on the output error. The contribution of the timing noise on the output error is compared with the ones of input and output noise in terms of bias and variance. It is shown that the influence of the timing noise on the output error variance is not significant. However, it is shown that the bias on the output error induced by the timing noise cannot be neglected for the first measurement time instants. Consequently, the timing noise may affect the time-constant estimate of a first-order transfer function. Indeed estimation results have pointed out that the uncertainty on the time constant estimate is much larger than for the static gain estimate. An application to *in vitro* data is developed in the second part of this paper. It is shown how the timing noise can be taken into account by bounded-error estimation algorithms based on interval analysis. The timing

noise is described as a bounded error and no further hypothesis about probability distributions is stated. The results presented herein emphasize the effectiveness of such an bounded-error estimation approach in such an experimental framework. However, in this study case, it is shown that the consequences of the output noise on the parameter estimates are more significant than the effects of the timing noise.

References

- [1] M. Barberi-Heyob, P.-O. Védérine, J.-L. Merlin, R. Millon, J. Abecassis, M.-F. Poupon, and F. Guillemin, “Wild-type p53 gene transfer into mutated p53 HT29 cells improves sensitivity to photodynamic therapy via induction of apoptosis,” *Int. J. Oncol.*, no. 24, pp. 951–958, 2004.
- [2] J. G. Moser, *Photodynamic Tumor Therapy: 2nd and 3rd Generation*. Gordon & Breach Science Publishers, 1998.
- [3] R. Bonnett, *Chemical Aspects of Photodynamic Therapy*. Gordon and Breach Science Publishers, 2000.
- [4] H. Barr and C. Hopper, *Photodynamic Therapy*. Royal Soc.of Medicine P., 2005.
- [5] M. P. Goldman, *Photodynamic Therapy*. W.B. Saunders Company, 2005.
- [6] N. D. Evans, R. J. Errington, M. J. Chapman, P. J. Smith, M. J. Chappell, and K. R. Godfrey, “Compartmental modelling of the uptake kinetics of the anti-cancer agent topotecan in human breast cancer cells,” *International Journal of Adaptive Control and Signal Processing*, vol. 19, pp. 395–417, 2005.
- [7] C. L. Beck, H.-H. Lin, and M. Bloom, “Multivariable modeling and control of anesthetic pharmacodynamics,” in *Proc of the CNRS-NSF Workshop on Biology and Control Theory: current challenges*, Toulouse, France, April 2006.
- [8] R. Ali, L. Campbell, N. D. Evans, R. J. Errington, K. R. Godfrey, P. J. Smith, and M. J.

Chappell, “A PK-PD model of cell cycle response to topotecan,” in *Proc of the IFAC Modelling and Control in Biomedical Systems*, Reims, France, September 2006, pp. 477–482.

- [9] J. Kusuma and V. K. Goyal, “Signal parameter estimation in the presence of timing noise,” in *Proc of the 40th Conference on Information Science and Systems, CISS’06*, Princeton, NJ, USA, March 2006.
- [10] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis*. Springer, 2001.
- [11] P. Eykhoff, *System Identification*. John Wiley & Sons, 1979.
- [12] L. Ljung, *System Identification: Theory For The User*. Englewood Cliffs, NJ: Prentice Hall, 1987.
- [13] E. Walter and L. Pronzato, *Identification of Parametric Models from experimental data*. Springer-Verlag, Masson, 1997.
- [14] T. Bastogne, L. Tirand, M. Barberi-Heyob, and A. Richard, “Modélisation système de la thérapie photodynamique,” in *4ème Conférence Internationale Francophone d’Automatique*, Bordeaux, France, May-June 2006.
- [15] K. Mahata and H. Garnier, “Identification of continuous-time errors-in-variables models,” *Automatica*, vol. 42, no. 9, pp. 1477–1490, September 2006.
- [16] T. Bastogne, L. Tirand, M. Barberi-Heyob, and A. Richard, “System identification of photosensitizer uptake kinetics in photodynamic therapy,” in *Proc of the 6th IFAC Symposium on Modelling and Control in Biomedical System (including biological systems)*, Reims, France, Sept. 2006.
- [17] I. Braems, N. Ramdani, M. Kieffer, L. Jaulin, E. Walter, and Y. Candau, “Guaranteed characterization of thermal conductivity and diffusivity in presence of model uncertainty,” accepted in *Inverse Problems in Science and Engineering*, 2006.

Lemma .1

$$\mathbb{E}\{e^{a(\frac{1}{2}-U)}\} = \frac{\sinh(\frac{a}{2})}{\frac{a}{2}}, \quad (25)$$

where a is a non-null constant and U is a random variable distributed according to a uniform law on $[0, 1]$.

A Proof of the Proposition 4.1

Since G_u and G_y^j are centered and since U_t^j and G_u are independent, then

$$\begin{aligned} \mathbb{E}\{e_y(t_j)\} &= k \cdot u_0 \cdot \left(1 - \mathbb{E}\{e^{-\frac{1}{T}(t_j - \frac{\tau}{2} + \tau U_t^j)}\}\right) \\ &\quad - k \cdot u_0 \cdot \left(1 - e^{-\frac{t_j}{T}}\right) \end{aligned} \quad (26)$$

$$= k \cdot u_0 \cdot e^{-\frac{t_j}{T}} \left(1 - \mathbb{E}\{e^{-\frac{1}{T}(-\frac{\tau}{2} + \tau U_t^j)}\}\right) \quad (27)$$

$$= k \cdot u_0 \cdot e^{-\frac{t_j}{T}} \left(1 - \mathbb{E}\{e^{\frac{\tau}{T}(\frac{1}{2}U_t^j)}\}\right). \quad (28)$$

It can be deduced from Lemma .1 that

$$\mathbb{E}\{e_y(t_j)\} = k \cdot u_0 \cdot e^{-\frac{t_j}{T}} \left(1 - \frac{\sinh(\frac{\tau}{2T})}{\frac{\tau}{2T}}\right). \quad (29)$$

B Proof of the Proposition 4.2

From (12), $e_y(t_j)$ is rewritten such that

$$e_y(t_j) = X_1 + X_2 - y_{\mathcal{M}}(t_j), \quad (30)$$

$$(31)$$

with

$$X_1 = k \cdot (u_0 + \sigma_u G_u) \cdot \left(1 - e^{-\frac{1}{T}(t_j - \frac{\tau}{2} + \tau U_t^j)}\right) \quad (32)$$

$$X_2 = \sigma_y G_y^j. \quad (33)$$

X_1 and X_2 are independent. It can be deduced that

$$\text{Var}(e_y(t_j)) = \text{Var}(X_1) + \text{Var}(X_2) \quad (34)$$

$$= \text{Var}(X_1) + \sigma_y^2. \quad (35)$$

Let us compute the expectation of X_1 .

$$\mathbb{E}\{X_1\} = k \cdot u_0 \cdot \left(1 - e^{-\frac{t_j}{T}} \mathbb{E}\{e^{\frac{\tau}{T}(\frac{1}{2}-U_t^j)}\}\right) \quad (36)$$

$$= k \cdot u_0 \cdot \left(1 - e^{-\frac{t_j}{T}} \frac{\sinh(\frac{\tau}{2T})}{\frac{\tau}{2T}}\right) \quad (37)$$

according to the Lemma .1. The expectation of X_1^2 is given by

$$\mathbb{E}\{X_1^2\} = k^2 \mathbb{E}\{(u_0 + \sigma_u G_u)^2\} \mathbb{E}\{X_3^2\} \quad (38)$$

$$= k^2 (u_0^2 + \sigma_u^2) \mathbb{E}\{X_3^2\} \quad (39)$$

where

$$X_3 = 1 - e^{-\frac{1}{T}(t_j - \frac{\tau}{2} + \tau U_i^j)} \quad (40)$$

$$\begin{aligned} \mathbb{E}\{X_3^2\} &= 1 + \mathbb{E}\{e^{-\frac{2}{T}(t_j - \frac{\tau}{2} + \tau U_i^j)}\} - 2\mathbb{E}\{e^{-\frac{1}{T}(t_j - \frac{\tau}{2} + \tau U_i^j)}\} \\ &= 1 + e^{-\frac{2t_j}{T}} \mathbb{E}\{e^{\frac{2\tau}{T}(\frac{1}{2} - U_i^j)}\} - 2e^{-\frac{t_j}{T}} \mathbb{E}\{e^{\frac{\tau}{T}(\frac{1}{2} - U_i^j)}\} \\ &= 1 + e^{-\frac{2t_j}{T}} \frac{\sinh(\frac{\tau}{T})}{\frac{\tau}{T}} - 2e^{-\frac{t_j}{T}} \frac{\sinh(\frac{\tau}{2T})}{\frac{\tau}{2T}}. \end{aligned} \quad (41)$$

We finally obtain

$$\begin{aligned} Var\{X_1\} &= k^2 (u_0^2 + \sigma_u^2) \left(1 + e^{-\frac{2t_j}{T}} \frac{\sinh(\frac{\tau}{T})}{\frac{\tau}{T}} \right. \\ &\quad \left. - 2e^{-\frac{t_j}{T}} \frac{\sinh(\frac{\tau}{2T})}{\frac{\tau}{2T}} \right) - k^2 u_0^2 \left(1 - e^{-\frac{t_j}{T}} \frac{\sinh(\frac{\tau}{2T})}{\frac{\tau}{2T}} \right)^2 \\ &= k^2 (u_0^2 + \sigma_u^2) \left(1 + e^{-\frac{2t_j}{T}} \frac{\sinh(\frac{\tau}{T})}{\frac{\tau}{T}} \right. \\ &\quad \left. - 2e^{-\frac{t_j}{T}} \frac{\sinh(\frac{\tau}{2T})}{\frac{\tau}{2T}} \right) \\ &\quad - k^2 u_0^2 \left(1 - 2e^{-\frac{t_j}{T}} \frac{\sinh(\frac{\tau}{2T})}{\frac{\tau}{2T}} + e^{-\frac{2t_j}{T}} \left(\frac{\sinh(\frac{\tau}{2T})}{\frac{\tau}{2T}} \right)^2 \right) \\ &= k^2 \left(\sigma_u^2 + (u_0^2 + \sigma_u^2) e^{-\frac{2t_j}{T}} \frac{\sinh(\frac{\tau}{T})}{\frac{\tau}{T}} \right. \\ &\quad \left. - 2\sigma_u^2 e^{-\frac{t_j}{T}} \frac{\sinh(\frac{\tau}{2T})}{\frac{\tau}{2T}} \right. \\ &\quad \left. - u_0^2 e^{-\frac{2t_j}{T}} \left(\frac{\sinh(\frac{\tau}{2T})}{\frac{\tau}{2T}} \right)^2 \right) \end{aligned} \quad (42)$$

Equation (4.2) is then a direct consequence of (35).

Traitement de signaux physiologiques

M. Caparos, V. Louis-Dorr, R. Ranta

La problématique scientifique commune aux différentes applications de ce projet est la compréhension et la caractérisation de certains mécanismes inhérents à des structures anatomo-fonctionnelles humaines à partir de signaux mono-dimensionnels, mono- ou multi-capteurs et évoluant dans le temps. Les chercheurs de ce projet tentent de développer des méthodes de traitement des signaux fonction de l'objectif et du contexte médical. Le domaine principal d'application est électroencéphalographie (EEG), et en particulier l'analyse des EEG épileptiques. D'autres travaux concernent l'analyse des sons abdominaux (sujet de recherche arrêté depuis peu) et l'analyse de mouvement (sujet lancé depuis un an et encore en phase d'évaluation).

Les compétences des membres du projet sont principalement l'analyse des signaux et leur pré-traitement. Cet article présente, dans une première section, une méthodologie d'analyse pour les EEG de surface. La deuxième partie est consacré aux techniques d'analyse et débruitage par ondelettes, appliquées déjà sur les sons abdominaux et pas encore testées sur les EEG.

1 Mesure de relations inter-structures en EEG

Les synchronisations électrophysiologiques entre différentes structures cérébrales sont à l'origine de certains fonctionnements neurologiques cérébraux. Cependant des décharges électrophysiologiques anormales et excessives produisent de graves troubles comme l'épilepsie [1]. L'analyse des signaux EEG est une des modalités qui aide à la compréhension et au diagnostic de ces mécanismes paroxystiques. En EEG de profondeur (SEEG) les relations statistiques existant entre deux signaux électrophysiologiques peuvent être mises en rapport avec les couplages fonctionnels entre certaines structures cérébrales sous-jacentes [2]. Pour l'épilepsie, connaître l'évolution des mesures de relation entre structures au départ d'une crise revient donc à comprendre et à modéliser les échanges inter-structures qui entrent en jeu lors de la mise en place d'une crise d'épilepsie [3, 1, 4]. La détermination des directions de propagation de l'activité EEG par la mesure de ces couplages est donc un facteur essentiel dans la compréhension de l'épilepsie [5]. La bibliographie fait état de deux grandes familles de méthodes permettant la mesure des synchronisations entre zones cérébrales au départ des crises. Les modélisations du signal fournis par des couples d'électrodes EEG par méthodes paramétriques (section 1.1) et non-paramétriques (section 1.2) sont décrites et analysées. L'originalité de ce travail est de déterminer si les synchronisations interstructures mises en évidence en profondeur au cours de la crise sont transférées en surface.

1.1 Mesures paramétriques

Dans le cadre de l'étude des séries temporelles non déterministes, il est souvent intéressant de décrire l'évolution des données par un modèle mathématique.

1.1.1 Les modèles AutoRégressifs

Présentation générale des modèles AR Le modèle mathématique le plus couramment utilisé pour le traitement de l'EEG est la fonction de transfert rationnelle dont la forme exacte est déterminée par l'estimation des coefficients devant ses *pôles* et ses *zéros*. Si la fonction de transfert ne contient que des *pôles*, le modèle est appelé *AutoRégressif (AR)*. Si il n'y a que des *zéros*, on parle de *moyenne mobile (MA, Moving Average)* [6]. Une combinaison de ces deux modèles est évidemment possible et dans ce cas, nous avons un modèle *Autorégressif à moyenne mobile, (communément appelé ARMA : Autoregressive with a Moving Average)*.

Pour pouvoir appliquer une modélisation autorégressive à un signal, une des hypothèses est que celui-ci soit stationnaire, *i.e.* que ses caractéristiques statistiques (les moments de tous les ordres, son contenu spectral, ...) ne varient pas avec le temps. Toutefois, quand on travaille sur des signaux physiques comme l'EEG, cette hypothèse n'est que très rarement vérifiée. Dans ce cas, l'approximation la plus courante est

de considérer des segments très courts dans lesquels l’hypothèse de quasi-stationnarité est vérifiée. Il faut alors estimer un modèle pour chacun de ces segments. Il existe plusieurs méthodes pour déterminer la longueur de ceux-ci. La plus simple est de la fixer arbitrairement en fonction de connaissance *a priori*, alors que les méthodes les plus sophistiquées mesurent continuellement les caractéristiques statistiques du signal et déterminent la fin du segment après détection d’un saut.

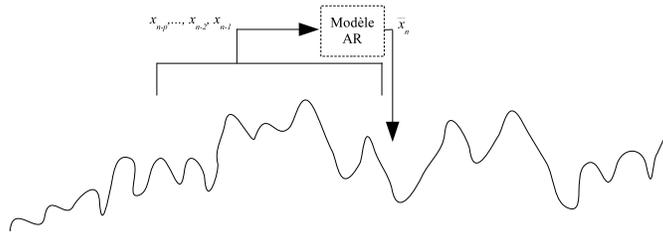


FIG. 1 – Principe de fonctionnement de l’estimation de l’échantillon suivant par un modèle AR.

Le principe de la modélisation AR peut être formulé aussi bien dans le domaine fréquentiel tel un problème d’identification spectrale que dans le domaine temporel tel un problème de prédiction linéaire.

Propagation du signal électrophysiologique Récemment, dans [7] un modèle AR multivariable d’ordre six, a été appliqué à une grandeur permettant de mesurer des relations causales entre différentes variables c’est le concept de “Granger causality” [8]. Ce dernier, qui a été développé au départ pour des applications économiques, a depuis été repris pour mesurer la propagation de l’activité lors des crises d’épilepsie et a été étendu aux modèles AR non-linéaires [9]. Le principal défaut des mesures décrites ci-dessus est que même si elles utilisent un modèle multivariable, elles ne permettent l’analyse des relations que deux par deux. Or dans [5], les auteurs démontrent que l’estimation des relations se basant uniquement sur deux canaux de l’EEG donnent souvent des résultats erronés. Pour modéliser correctement le flux d’informations entre les voies de l’EEG, au cours d’une crise d’épilepsie, il est impératif d’utiliser une mesure contenant l’information globale. Une telle méthode a été introduite par Kaminski et Blinowska [10]. Celle-ci a depuis été reprise par l’équipe de Franaszczuk [11, 12, 13] et améliorée.

Identification d’un modèle AR Multivariable Pour un modèle AR multivariable, comme pour un modèle scalaire, l’élément courant $x(k)$ est considéré comme une combinaison linéaire d’un certain nombre d’observations passées du processus, auxquelles vient s’ajouter une entrée aléatoire non-corrélée [14].

$$\mathbf{x}(k) = -\mathbf{A}(1)\mathbf{x}(k-1) - \mathbf{A}(2)\mathbf{x}(k-2) + \dots - \mathbf{A}(p)\mathbf{x}(k-p) + \mathbf{e}(k) \quad (1)$$

avec $x_i(k), i = 1 \dots d$, l’observation de la voie i à l’instant k et $\mathbf{e}(k)$ un vecteur bruit non-corrélé de dimension d avec une matrice de variance-covariance \mathbf{V}_e . Les $\mathbf{A}(i), i = 1, \dots, p$, sont des matrices de dimension $d \times d$, les coefficients du modèle AR qui sont estimés à partir des échantillons $\mathbf{x}(k), k = 1, \dots, N, N$ le nombre de points du signal et p l’ordre du modèle AR. L’estimation de l’échantillon courant $\tilde{\mathbf{x}}(k)$ s’écrit :

$$\tilde{\mathbf{x}}(k) = -\sum_{i=1}^p \mathbf{A}(i)\mathbf{x}(k-i) \quad (2)$$

L’erreur de prédiction \mathbf{e} , appelée aussi résidu de l’estimation, est alors donnée par :

$$\mathbf{e}(k) = \mathbf{x}(k) - \tilde{\mathbf{x}}(k) \quad (3)$$

Le calcul des coefficients du modèle AR revient donc au problème de la minimisation de l’erreur de prédiction par la détermination des coefficients $\mathbf{A}(i)$. Cela peut être fait par la méthode des moindres carrés en déterminant les coefficients $\mathbf{A}(i)$ qui minimisent le critère $J(n)$ suivant :

$$J(n) = \sum_{k=p+1}^n \mathbf{e}^T(k)\mathbf{e}(k) \quad (4)$$

Ce critère peut être minimisé à partir des dérivées partielles de $J(n)$ par rapport aux éléments des matrices $\mathbf{A}(i)$.

De nombreuses méthodes d'estimation de l'ordre existent dans la littérature appliquée à l'EEG. Dans [15], il a été montré que le plus satisfaisant des critères est celui décrit par Akaike en 1974 [16]. Il est appelé critère AIC (Akaike's Information Criterion) et est fondé sur l'estimation du maximum de vraisemblance de la fonction de densité de probabilité du signal. L'ordre qui convient le mieux à la modélisation est celui qui minimise le critère AIC :

$$AIC(p) = N \ln(\tilde{\sigma}_e^2(p)) + 2p, \quad (5)$$

où $\tilde{\sigma}_e^2(p)$ est la variance de l'erreur obtenue pour un modèle d'ordre p .

La détermination de l'ordre dépend *a priori* souvent aussi de l'application, du type de signaux à modéliser et de l'utilisation visée par le modèle. Ainsi, quand il s'agit de traiter l'EEG de patients épileptiques, certains auteurs utilisent des modèles d'ordre très petit : 3 dans [12] ou 4 dans [17]. D'après [18], il faut plutôt un ordre 6 pour modéliser correctement le signal.

Mesures paramétriques de relations inter-structures Deux grandeurs basées sur une modélisation paramétrique permettent la mesure de l'interdépendance entre deux structures cérébrales, elles ont été appliquées aux signaux de profondeur (SEEG ou ECoG) mais peu d'études ont été faites en les appliquant aux EEG de surface.

La cohérence n'a pas été développée spécifiquement pour être utilisée sur les signaux EEG. Toutefois, elle est proche d'une mesure de corrélation linéaire entre plusieurs voies d'un signal et donc la mesure de relations inter-structures pour les EEG en est une application particulièrement intéressante. Pour la calculer, on peut réécrire l'équation (1) dans le domaine fréquentiel, ce qui donne :

$$\mathbf{X}(f) = \mathbf{H}(f)\mathbf{E}(f), \quad (6)$$

où $\mathbf{X}(f)$ est le vecteur contenant les transformées de Fourier X_i des voies x_i :

$$X_i(f) = \sum_{k=0}^{N-1} x_i(k) e^{-2j\pi f \frac{k}{N}}$$

Après avoir vérifié que les matrices $\mathbf{A}(i)$ sont inversibles, on obtient les coefficients de la matrice $\mathbf{H}(f)$:

$$\mathbf{H}_{lm}(f) = \left(\sum_{k=0}^p \mathbf{A}_{lm}(k) e^{-2j\pi f \frac{k}{N}} \right)^{-1} \quad (7)$$

$\mathbf{H}_{lm}(f)$ est le vecteur représentant l'évolution du contenu de la $l^{\text{ième}}$ ligne et de la $m^{\text{ième}}$ colonne de $\mathbf{H}(f)$. $\mathbf{H}(f)$ est appelé matrice de transfert. Rappelons que $\mathbf{A}(0)$ est la matrice identité. On peut alors construire la matrice de densité spectrale [19] :

$$\mathbf{S}(f) = \mathbf{H}(f)\mathbf{V}_e\mathbf{H}(f)^* \quad (8)$$

où \mathbf{V}_e est la matrice de variance-covariance du bruit de modélisation \mathbf{e} et $\mathbf{H}(f)^*$ est la transposée de la matrice conjuguée de $\mathbf{H}(f)$. $\mathbf{S}(f)$ est couramment utilisée dans la plupart des applications d'analyse spectrale multivariable car elle permet de calculer différentes caractéristiques spectrales d'un signal comme la puissance, la cohérence ainsi qu'un ensemble de grandeurs dérivées de celle-ci.

La cohérence, proprement dite, est définie comme le rapport de l'inter-spectre au carré sur le produit des auto-spectres d'un signal, on a donc :

$$C_{lm}(f) = \frac{|\mathbf{S}_{lm}(f)|^2}{\mathbf{S}_{ll}(f)\mathbf{S}_{mm}(f)} \quad (9)$$

Où \mathbf{S}_{lm} est l'élément (l, m) de la matrice spectrale. Les valeurs de la cohérence sont normalisées entre 0 et 1 et permettent la mesure de la dépendance linéaire entre $\mathbf{x}_l(t)$ et $\mathbf{x}_m(t)$ qui sont les $l^{\text{ième}}$ et $m^{\text{ième}}$ éléments du vecteur \mathbf{x} pris à l'instant t . La cohérence est calculée pour chaque fréquence f . Si sa valeur est proche de 1 pour une fréquence donnée, les deux signaux ont une interdépendance maximale à cette fréquence. A l'opposé, une valeur proche de 0 indique l'indépendance de ces processus à la fréquence considérée.

Le principal défaut de la cohérence, pour l'application qui nous intéresse, est qu'elle est symétrique, ie $C_{lm}(f) = C_{ml}(f)$, pour toutes les fréquences f .

Pour remédier à ce défaut, Kaminski et Blinowska (1991) [10] ont introduit la *Directed Transfer Function* (DTF), elle donne une indication sur le sens du flux d'informations [20].

Elle est définie directement à partir de la matrice de transfert du système $\mathbf{H}(f)$ par l'équation :

$$DTF_{lm}(f) = \frac{|\mathbf{H}_{lm}(f)|^2}{\sum_{k=1}^d |\mathbf{H}_{lk}(f)|^2} \quad (10)$$

1.2 Mesures non-paramétriques

L'hypothèse de stationnarité n'est pas toujours vérifiée, il est donc nécessaire de comparer les résultats obtenus par l'analyse de la cohérence et la DTF à d'autres grandeurs n'utilisant pas de modélisation. Ces deux grandeurs sont la corrélation linéaire et la corrélation non-linéaire entre les signaux.

1.2.1 Corrélation linéaire en sous-bandes

La formule de calcul de la corrélation est bien connue et est donnée par :

$$r_{ij} = \frac{cov(\mathbf{x}_i, \mathbf{x}_j)}{\sqrt{var(\mathbf{x}_i)var(\mathbf{x}_j)}} \quad (11)$$

où \mathbf{x}_i et \mathbf{x}_j sont deux voies de l'EEG.

Les valeurs prises par r_{ij} sont comprises entre -1 et 1 . Plus la corrélation est proche de 1 , en valeur absolue, plus la relation entre les deux structures est grande. En effet, suivant la position des capteurs mesurant l'activité électrique deux signaux très liés peuvent être en phase ou en anti-phase. Une valeur absolue de la corrélation proche de 1 indique que le degré de relation entre les deux suites temporelles est fort. Une valeur proche de 0 indique, au contraire, un faible niveau de corrélation. Pour l'utilisation qui en est faite dans le cas présent, il est intéressant de bien différencier fortes et faibles relations statistiques, nous avons donc choisi d'élever la relation au carré et d'utiliser l'indice :

$$r_{ij}^2 = \frac{(cov(\mathbf{x}_i, \mathbf{x}_j))^2}{var(\mathbf{x}_i)var(\mathbf{x}_j)} \quad (12)$$

r_{ij}^2 prend ses valeurs entre 0 et 1 . Comme la cohérence, r_{ij}^2 est une grandeur symétrique et donc $r_{ij}^2 = r_{ji}^2$. Cet indice est calculé sur une fenêtre glissante de longueur d , définie expérimentalement à partir des connaissances apportées par les neurologues. Ceci permet de suivre l'évolution de l'indice au cours du temps et de mettre en évidence la dynamique de la crise.

Outre, le caractère symétrique, le principal défaut de cette méthode est l'obligation d'utiliser un indice par bande de fréquences pour décrire globalement le développement d'une crise. Ce problème n'est pas insurmontable, mais rend plus complexe le choix des différents paramètres utilisés dans la méthodologie en imposant un ajustement pour chacun d'entre eux. Comme un certain nombre de publications [21, 22, 9] montrent que les structures cérébrales formeraient plutôt un réseau non-linéaire lors d'une crise d'épilepsie, nous avons étudié l'application d'un coefficient de régression non-linéaire aux EEG.

1.2.2 Coefficient de régression non-linéaire h_{XY}^2

Son but principal est de déterminer une fonction univoque permettant de relier deux variables distinctes pour lesquelles on fait l'hypothèse qu'il existe une relation de dépendance fonctionnelle. La détermination de cette dépendance consiste à estimer, au sens des moindres carrés, une courbe de régression non-linéaire.

Dans la réalité, la dépendance fonctionnelle n'est pas analytiquement évidente, mais on dispose à la fois d'un certain nombre d'observations et de connaissances physiques grâce à l'étude des signaux SEEG sur la nature de la relation recherchée [2]. Les données disponibles se présentent sous la forme de couples d'observations qui dans notre cas sont les enregistrements de deux voies \mathbf{x}_i et \mathbf{x}_j de l'EEG. Afin d'alléger la notation, pour la suite de cette partie on note X (resp. Y) les voies notées précédemment \mathbf{x}_i (resp. \mathbf{x}_j). Dans ce cas, x_k (resp. y_l) représentent les valeurs des deux voies X et Y aux instants k et l .

L'intensité de la liaison entre deux distributions statistiques est mesurée par le rapport de détermination (ou de corrélation). Il est défini comme le rapport de la variance inter-classe (expliquée) sur la variance totale (intra-classe ou marginale). Celui-ci exprime donc la proportion de la variance marginale représentée

(ou expliquée) par la variance des moyennes conditionnelles. Le rapport de détermination de Y sachant X s'écrit donc :

$$h_{XY}^2 = \frac{\text{var}_{expl}(Y)}{\text{var}(Y)} = \frac{\text{var}(\overline{y_k})}{\text{var}(Y)}, \quad (13)$$

avec $\overline{y_k}$ la moyenne conditionnelle de Y sachant x_k .

La variance marginale peut s'écrire à partir de la variance des moyennes conditionnelles et de la moyenne des variances conditionnelles sous la forme :

$$\text{var}(Y) = \text{var}(\overline{y_k}) + \overline{\text{var}_c(Y)} \quad (14)$$

En remplaçant (14) dans (13) et en simplifiant on obtient :

$$h_{XY}^2 = 1 - \frac{\overline{\text{var}_c(Y)}}{\text{var}(Y)} \quad (15)$$

Dans le cadre de l'étude de l'EEG, ce coefficient est utilisé pour estimer un degré de couplage entre plusieurs structures cérébrales. Or, comme une certaine distance sépare ces mêmes structures, l'information met un certain temps τ pour se propager entre elles. Ce temps de propagation τ peut être estimé en maximisant le coefficient h_{XY}^2 en fonction des différents décalages temporels τ . En pratique, on réitère le calcul de $h_{XY}^2(\tau)$ entre le signal X et Y_τ (versions décalées de Y) en faisant varier τ dans un intervalle défini par les médecins. Seule, la valeur maximale $h_{XY,max}^2$ de $h_{XY}^2(\tau)$ est conservée.

Ce coefficient est appelé, dans la suite, coefficient de corrélation non-linéaire, car la variance conditionnelle $\overline{\text{var}_c(Y)}$ est estimée en considérant une régression non-linéaire (plus précisément, linéaire par morceaux).

Une remarque importante est que les variables X et Y peuvent avoir un "grand" coefficient de corrélation non-linéaire h_{YX}^2 (ou h_{XY}^2) en n'étant pourtant pas linéairement corrélés. En effet, la seule relation liant ces deux types de corrélation est la suivante :

$$0 \leq r_{XY}^2 = r_{YX}^2 \leq \min(h_{XY}^2, h_{YX}^2) \leq 1 \quad (16)$$

Estimation des paramètres optimaux pour $h_{XY,max}^2$ Pour décrire correctement l'évolution des signaux EEG, il faut étudier l'évolution des coefficients de corrélation h_{XY}^2 dans le temps. Nous avons identifié quatre paramètres utilisateurs qui interviennent dans la mise en oeuvre de ce calcul :

1. Une fenêtre de calcul de d secondes ($d * f_e$ points, ou f_e est la fréquence d'échantillonnage),
2. Un pas de calcul p_{as} , pour régler l'intervalle entre deux fenêtres successives,
3. Un certain nombre d'intervalles n_{cel} permettant le calcul des droites affines qui donnent l'approximation de la régression non-linéaire,
4. Le degré de liberté τ_{max} laissé dans le calcul de la régression non-linéaire et qui permet d'optimiser le coefficient pour le retard existant entre les deux signaux X et Y , en maximisant h_{XY}^2 .

Le réglage de ces quatre paramètres s'est fait par essais successifs car il n'existe pas de méthode établie. On peut toutefois se baser sur des notions physiologiques afin de déterminer *a posteriori* si les résultats obtenus concordent avec ceux attendus.

1.3 Comparaison entre les méthodes

La figure 2 présente un exemple comparatif de l'évolution de l'estimation des synchronisations sur des signaux EEG réels. Pour quantifier la qualité du résultat, nous avons choisi d'utiliser plusieurs indices correspondants à la moyenne et à l'écart-type de l'évolution des corrélations. Ces indices sont calculés sur des fenêtres de 50 secondes pendant la crise et longtemps avant la crise. Le rapport des moyennes donne la méthode pour laquelle l'augmentation de la synchronisation est la plus significative entre une fenêtre prise dans la crise et une fenêtre en période inter-critique. Le calcul de ces indices a été appliqué à l'ensemble des 43 patients pour chacune des 87 crises. Les résultats moyens obtenus pour chaque méthode sont présentés dans le tableau 1.

L'analyse de ces indices, révèle plusieurs points intéressants :

1. Les moyennes n'évoluent pas significativement entre la DTF et le h^2 , que ce soit pendant la période critique ou la période inter-critique. Cela implique donc bien entendu que le rapport entre les moyennes reste le même,

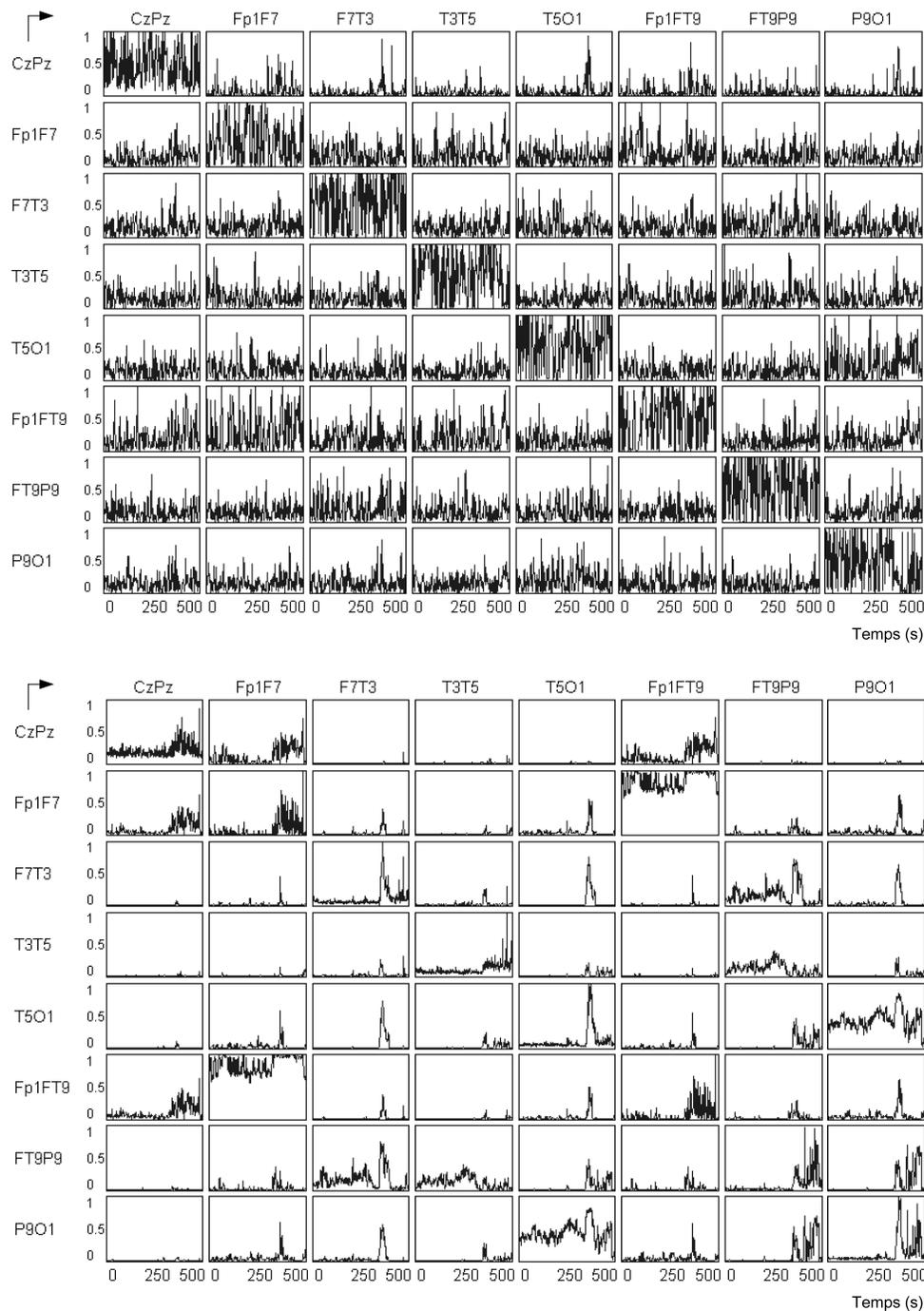


FIG. 2 – Exemple de mesure d'évolution des coefficients de la DTF (en haut) et de la régression non-linéaire h^2 (en bas) entre les huit voies d'un signal EEG réel choisies par le neurologue.

2. Les moyennes mesurées pour les méthodes basées sur les corrélations en sous bandes (delta, theta et alpha) sont plus élevées. Le problème est que les moyennes mesurées sur les périodes intercritiques sont du même ordre que celles mesurées pendant la crise. Ces résultats ne correspondent pas à ce que nous attendions,
3. Globalement les écarts-types hors crises sont plus faibles que pendant la crise.
4. Les écarts-type les plus faibles sont observés pour la méthode utilisant une régression non-linéaire.

	moyenne pendant crise	moyenne hors crise	écart-type pendant crise	écart-type hors crise
DTF : 0.5-30 Hz	0.22	0.15	0.16	0.11
h^2 : 0.5-30 Hz	0.23	0.16	0.09	0.06
r^2 : 0.5-30 Hz	0.22	0.27	0.13	0.14
r^2 : <i>delta</i>	0.25	0.26	0.16	0.14
r^2 : <i>thta</i>	0.29	0.28	0.19	0.18
r^2 : <i>alpha</i>	0.25	0.23	0.16	0.12

TAB. 1 – Indices caractérisant l'évolution du niveau de relation pour l'ensemble des 87 crises des 43 patients sélectionnés entre les périodes critiques et inter-critiques. Le choix des patients est détaillé dans la première partie du chapitre suivant.

1.3.1 Conclusions

A la vue des résultats obtenus sur ces indices, le choix de l'algorithme de mesure de relation n'est donc pas une chose aisée. L'application sur signaux simulés a montré que la méthode basée sur la cohérence peut être facilement éliminée pour les raisons suivantes : elle ne permet pas la discrimination du sens d'un transfert d'énergie et si les signaux sont trop proches, elle sépare difficilement les "vraies" relations des "fausses". Il n'y a qu'une analyse de l'importance relative d'une voie par rapport à toutes les autres qui peut donner cette information.

En simulation, la DTF donne de très bons résultats car l'influence d'une voie sur une autre est très visible et seules les "vraies" relations apparaissent. Cette méthode souffre cependant d'un problème de fond quand elle est appliquée à des EEG réels. Elle s'appuie sur l'identification d'un modèle stationnaire sur des signaux non-stationnaires par essence. Nous avons vu que sur 5 secondes, l'hypothèse de pseudo-stationnarité est généralement vraie pour les EEG, ce qui justifie tout de même l'emploi de ce type de méthode. Pour identifier un modèle, il faut choisir un ordre suffisant pour décrire le signal. Dans la bibliographie, on ne trouve pas de consensus quant au choix de l'ordre. D'après nos propres estimations sur le critère AIC, on trouve qu'un ordre six peut convenir, bien que le critère n'admette pas réellement de minimum. Cela est dommageable pour la qualité de la modélisation, car prendre un ordre inférieur à celui donné par le critère implique forcément une erreur de modélisation. Malgré ce choix extrêmement discutable et arbitraire, on observe que la DTF donne des résultats qui semblent être corrects.

Il faut ensuite départager les méthodes non-paramétriques. A la vue des résultats obtenus dans le tableau 1, il semble immédiat d'éliminer les méthodes en bandes de fréquence, car la discrimination entre signal critique et non-critique ne peut pas se faire aisément. De nombreuses publications [21, 22, 9] montrent que les structures cérébrales s'inter-connectent entre elle de manière plutôt non-linéaire pendant une crise. Cet un indice est défini à partir d'une régression non-linéaire dans l'ensemble de la bande fréquentielle d'intérêt de l'EEG de surface. En effet, les résultats obtenus montrent que la régression non-linéaire est plus performante pour l'application aux signaux EEG.

En considérant les différents points cités ci-dessus, le choix se restreint donc à la DTF, si on opte pour une méthodologie à base de modélisation paramétrique du signal EEG ou à une méthode utilisant la régression non-linéaire. Elles semblent avoir des performances quantitatives similaires en terme de différence de niveau de relation entre critique et inter-critique. Il semble toutefois que la variance de la DTF soit nettement plus importante, ce qui rend la détection d'événement moins précise. De plus la complexité de la méthodologie employée est nettement plus grande ainsi que le temps de calcul. Pour toutes ces raisons, nous avons choisi de ne conserver que la régression non-linéaire pour le reste de ce travail de recherche : c'est à dire l'estimation de la latéralisation (droite ou gauche) et la classification des crises d'épilepsie du lobe temporal en fonction de la localisation de l'origine des crises (mésio-temporale ou basale). La méthodologie et les résultats de ces travaux sont publiés dans différents congrès [23].

Enfin, un point important mais qui n'a pas été abordé pour l'instant est le pré-traitement des EEG avant toute mesure de synchronisation. Deux approches complémentaires nous semblent envisageables :

- le débruitage des signaux mesurés afin d'éliminer le bruit (quasi-) stationnaire qui les perturbe, sans pour autant détériorer l'information contenue. Une possibilité intéressante est le débruitage par ondelettes, décrit dans la section suivante de cet article.
- l'élimination des artefacts (d'origine oculaire, musculaire, etc.).

2 Débruitage par ondelettes

2.1 Décompositions atomiques. Ondelettes

Cette section commence avec la présentation d'une des principales représentations temps-fréquence – les décompositions atomiques.

L'idée fondamentale des représentations atomiques est qu'on peut écrire un signal $x(t)$ comme une somme de signaux élémentaires, les atomes temps-fréquence $a_{\tau,\xi}(t)$, localisés dans une fenêtre temporelle de largeur T autour de l'instant t et dans une bande fréquentielle de largeur B autour d'une fréquence ν :

$$x(t) = \sum_{t,\nu} X_{\tau,\xi} \cdot a_{\tau,\xi}(t), \quad (17)$$

où les $X_{\tau,\xi}$ sont des coefficients de pondération et τ et ξ représentent respectivement la fenêtre temporelle (autour de l'instant t) et la bande fréquentielle (autour de la fréquence ν) dans lesquelles l'atome $a_{\tau,\xi}(t)$ est localisé.

Dans le cas général, si t et ν sont des variables continues, les pavés temps-fréquence seront superposés, la somme de l'équation (17) deviendra une intégrale continue et la représentation sera redondante. Cependant, un des intérêts majeurs des décompositions atomiques est la possibilité de construire des bases, notamment des bases orthonormées. L'équation (17) peut être vue comme la représentation d'un vecteur (signal) par la somme de ses projections sur les vecteurs de base (les atomes temps-fréquence), les coefficients $X_{\tau,\xi}$ étant les valeurs scalaires de ces projections. C'est notamment le cas des décompositions en ondelettes.

On appelle ondelette mère une fonction $\psi(t)$ dans \mathbf{L}^2 qui a un certain nombre de caractéristiques :

– moyenne nulle :

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 ; \quad (18)$$

– norme (\mathbf{L}^2) unitaire :

$$\|\psi(t)\|_2 = 1 ; \quad (19)$$

Par dilatations et translations, on obtient une famille d'ondelettes :

$$\psi_{s,u} = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right), \quad (20)$$

où s est l'échelle et u la position temporelle. Si la fonction ψ est régulière (elle est dérivable p fois), sa transformée de Fourier $\hat{\psi}(f)$ tend vers 0 quand f tend vers ∞ . Comme pour $f = 0$ la transformée de Fourier est nulle (directement de (18)), l'ondelette est associée à un filtre passe-bande et permet une localisation en temps et en fréquence (la limite est fixée par le principe d'incertitude de Heisenberg). Une décomposition en ondelettes discrète est obtenue en considérant des dilatations s et des translations u discrètes.

Si on considère une suite d'échelles s_j finie, $s_j \leq s_M$, on associe à la famille d'ondelettes (20) une famille de fonctions d'échelle, obtenue par translation d'une fonction $\phi(t)$ de norme \mathbf{L}^2 unitaire, qui est l'agrégation des ondelettes d'échelle supérieure à s_M [24]. Si les ondelettes sont associées aux filtres passe-bande correspondant aux échelles, la fonction d'échelle est associée à un filtre passe-bas.

La décomposition discrète la plus courante est la décomposition dyadique [24], pour laquelle le pas de discrétisation fréquentiel est l'octave ($s = 2^j$, avec $j = 1 \dots M$). Le pas temporel p est lié à l'échelle (il varie entre 1 et $N/2^j$). La base de projection est donc constituée d'ondelettes $\psi_{j,p}$ et de fonctions d'échelle $\phi_{M,p}$, où M est la profondeur de la décomposition en ondelettes. Les coefficients de la décomposition d'un signal x sur cette base sont notés $w^{j,p}$, et le signal s'écrit comme :

$$x(t) = \sum_{j,p} w^{j,p} \psi_{j,p}(t) + \sum_p w^{M,p} \phi_{M,p}(t). \quad (21)$$

Les coefficients de pondération sont donnés par la transformée en ondelettes, qui est la projection de x sur les ondelettes et les fonctions d'échelle :

$$\begin{aligned} w^{j,p} &= \langle x, \psi_{j,p} \rangle = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{2^j}} \psi\left(\frac{t-p}{2^j}\right) dt, \\ w^{M,p} &= \langle x, \phi_{M,p} \rangle = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{2^M}} \phi\left(\frac{t-p}{2^M}\right) dt. \end{aligned} \quad (22)$$

2.2 Détection, estimation et débruitage par ondelettes

Une des applications les plus courantes des ondelettes est l'extraction du signal informatif c du signal mesuré x . Dans nos travaux, nous avons considéré un modèle discretisé classique ([25, 26]) du signal mesuré x : un mélange additif d'un signal informatif c et d'un bruit de mesure r :

$$x(k) = c(k) + r(k). \quad (23)$$

La première étape de traitement est la décomposition de ce signal sur une base orthogonale d'ondelettes. Le choix de la base de projection est généralement fait par l'utilisateur, mais peut être aussi le résultat d'un algorithme de recherche de la meilleure base, comme proposé par Coifman et Wickerhauser dans [27].

D'après la définition générale de Coifman et Wickerhauser [27], le débruitage par ondelettes revient à "l'extraction d'une structure cohérente" c du signal traité, ce qui, inversement, revient à considérer le bruit r comme non cohérent par rapport à la base d'ondelettes choisie, donc non-corrélé avec les fonctions de base. En fait, les segments du signal peu corrélés avec la base ont des coefficients $w^{j,p}$ faibles, et sont attribués au bruit. Par un seuil adapté sur le vecteur des coefficients, on peut séparer le bruit (partie incohérente) du signal (partie cohérente). Le débruitage par ondelettes est donc implémenté par seuillage (filtrage non-linéaire ou régression non-linéaire, selon Antoniadis *et al.* [25]) : les coefficients d'ondelette supérieurs à un seuil S sont considérés comme faisant partie du signal informatif et les faibles sont attribués au bruit.

Nous allons présenter dans la suite les principales techniques de seuillage et d'estimation de la valeur du seuil.

2.2.1 Calcul du seuil

Comme dans toute approche de type seuillage, le problème principal réside dans le calcul du seuil S . Un passage en revue très complet des méthodes de calcul est présenté par Antoniadis *et al.* dans [25]. Ces méthodes peuvent être partagées en deux classes, selon le calcul du seuil : directes et itératives.

Méthodes directes Les méthodes directes se déroulent en général en deux étapes :

1. Calcul du seuil (ou des seuils, échelle par échelle) à partir de la distribution des coefficients d'ondelette.
2. Seuillage et reconstruction du signal débruité à partir des coefficients significatifs.

En fait, les méthodes les plus connues de débruitage par ondelettes sont celles introduites par Donoho et Johnstone [28, 26] : *VisuShrink* et *SureShrink*. La première propose un seuil *universel* S_u déterminé à partir de l'énergie estimée du bruit. Cette méthode utilise l'hypothèse d'un bruit blanc gaussien superposé au signal. Le seuil S_u est égal à $\sigma\sqrt{2\log_e N}$, où N désigne le nombre de points du signal et σ^2 la variance du bruit. Puisque le bruit est considéré blanc, son énergie est équi-répartie sur toutes les bandes de fréquence (les échelles) de la décomposition. En conséquence, l'écart-type est estimé dans une bande où le signal informatif est considéré inexistant ou quasi-inexistant, notamment dans les plus hautes fréquences. Les auteurs utilisent un estimateur robuste de σ à partir de la valeur médiane $\theta_{w^{1,p}}$ des coefficients d'ondelette de la première échelle :

$$\hat{\sigma} = \frac{\theta_{w^{1,p}}}{0,6745}. \quad (24)$$

Cette estimation est appliquée, si la distribution du bruit est gaussienne, pour éviter le biais dû aux grandes valeurs (les coefficients du signal informatif) de l'estimateur classique :

$$\sigma = \sqrt{\frac{1}{N_1} \sum_{p=1}^{N_1} (w^{1,p})^2}, \quad (25)$$

où $N_1 = N/2$ est le nombre de coefficients d'ondelettes de l'échelle haute-fréquence.

Dans [24, 29] et en faisant appel à la théorie statistique des valeurs extrêmes, on montre que ce seuillage est asymptotiquement optimal : la probabilité d'avoir des coefficients supérieurs à $S = \sigma\sqrt{2\log_e N}$ tend vers 0 quand N tend vers ∞ .

Dans leur article de synthèse, Antoniadis *et al.* [25] présentent aussi d'autres améliorations des méthodes de calcul du seuil décrites dans la littérature : l'invariance par rapport aux translations du signal, la validation croisée etc., qui ne seront pas détaillées ici.

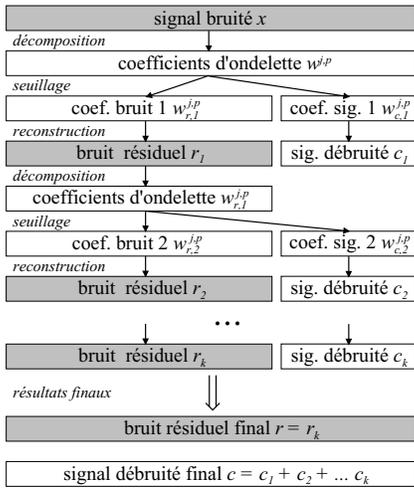


FIG. 3 – Schéma de principe d'un algorithme itératif de débruitage, selon Coifman et Wickerhauser [27]. A chaque itération, on annule les coefficients éliminés par seuillage, avant de reconstruire le bruit résiduel et le signal débruité.

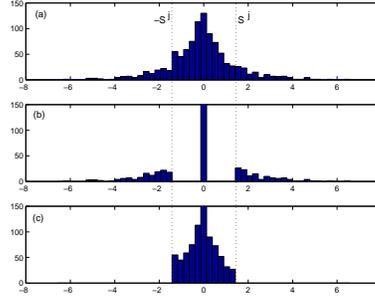


FIG. 4 – Histogramme des coefficients d'ondelette à l'échelle j (a) de l'estimation courante du bruit résiduel à l'itération l $\Omega_{r,l}^j$, (b) du signal informatif $\Omega_{\Delta c,l+1}^j$, (c) de la nouvelle estimation du bruit résiduel $\Omega_{r,l+1}^j$. Pour cet exemple, les coefficients d'ondelette du signal simulé suivent une loi de Laplace (l'axe des ordonnées n'est pas complètement représenté en (b) et (c)).

Méthodes itératives Les méthodes itératives de calcul du seuil ont été introduites indépendamment par plusieurs auteurs, dans différents domaines d'application. Le déroulement général de ces méthodes est le suivant :

1. Initialisation : quelles que soient l'échelle j et la position p , les coefficients du signal informatif $\{w_{c,0}^{j,p}\} = 0$ et les coefficients du bruit sont égaux aux coefficients du signal mesuré $\{w_{r,0}^{j,p}\} = \{w_x^{j,p}\}$.
2. Estimation du seuil : pour l'itération l , le seuil est calculé par différentes méthodes à partir de l'estimation courante du bruit r_l ou de ses coefficients d'ondelette $\{w_{r,l}^{j,p}\}$.
3. Seuillage : les coefficients significatifs (supérieurs au seuil) de l'estimation courante du bruit $\{w_{r,l}^{j,p}\}$ sont attribués au signal informatif, les autres étant attribués à la nouvelle estimation du bruit r_{l+1} .
4. Si un critère donné d'arrêt n'est pas atteint, retour au pas 2.

Un des premiers algorithmes est celui de Starck *et al.* [30], développé pour le traitement des images astronomiques. L'algorithme WTST-NST (*Wavelet Transform-based Stationary-Nonstationary Filter*) de Hadjileontiadis *et al.* [31] est très proche de celui de Starck, l'application étant médicale (sons pulmonaires et abdominaux). La différence entre les deux algorithmes réside dans le calcul de l'écart-type – robuste pour Starck (24) et classique pour Hadjileontiadis (25).

2.3 Algorithme de débruitage de Hadjileontiadis *et al.*

Considérons le modèle de l'équation (23), $\mathbf{x} = \mathbf{c} + \mathbf{r}$, où \mathbf{x} , \mathbf{c} et \mathbf{r} sont des vecteurs $1 \times N$. Les méthodes itératives de débruitage réécrivent cette équation comme $\mathbf{x} = \mathbf{c}_l + \mathbf{r}_l$, où l est l'indice d'itération. L'estimation courante du bruit \mathbf{r}_l (initialisée à $\mathbf{r}_0 = \mathbf{x}$) est décomposée sur une base d'ondelettes orthogonale :

$$\mathbf{r}_l = \sum_{j,p} w_{r,l}^{j,p} \psi^{j,p} + \sum_p w_{r,l}^{M,p} \phi^{M,p}, \quad (26)$$

où j est l'échelle, p l'indice de position temporelle et M la profondeur de la décomposition.

Notons $\Omega_{r,l}^j$ le vecteur des coefficients du bruit $\{w_{r,l}^{j,p}\}$ à l'échelle j et $\Omega_{r,l}$ le vecteur complet des coefficients du bruit. Par seuillage de $\Omega_{r,l}$, on obtient $\Omega_{\Delta c,l+1}^j$, le vecteur des grands coefficients qui sont considérés comme appartenant au signal informatif (voir figure 4). Le vecteur des coefficients du bruit $\Omega_{r,l+1}$, obtenu à partir de l'équation $\Omega_{\Delta c,l+1}^j + \Omega_{r,l+1} = \Omega_{r,l}$, est utilisé pour la reconstruction de l'estimation courante du bruit résiduel \mathbf{r}_{l+1} , qui sert à calculer le critère d'arrêt :

$$STC_{l+1} = \|\mathbf{r}_l\|^2 - \|\mathbf{r}_{l+1}\|^2 < \varepsilon, \quad (27)$$

où ε est un premier paramètre choisi par l'utilisateur. Le vecteur des coefficients informatifs $\Omega_{\Delta c,l+1}^j$ contribue à la reconstruction de \mathbf{c}_{l+1} , qui sert à calculer le signal informatif final (figure 3).

Un deuxième paramètre utilisateur F_a intervient dans le calcul du seuil S . L'algorithme WTST-NST de Hadjileontiadis *et al.* [31] calcule ce seuil en fonction de l'écart-type σ_l^j des coefficients d'ondelette $\{w_{r,l}^{j,p}\}$ du bruit estimé à chaque échelle. Pour chaque échelle, le seuil à l'itération l est recalculé comme :

$$S_{l+1}^j = F_a \cdot \sigma_l^j, \quad (28)$$

où l'expression de la constante $F_a = 3$ est validée par expertise médicale¹.

Schématiquement, l'algorithme de Hadjileontiadis *et al.* [31] s'écrit :

1. calculer σ_l^j comme $(\sigma_l^j)^2 = \frac{1}{N^j} \|\Omega_{n,l}^j\|^2$, où $N^j = \text{card}(\Omega_{n,l}^j)$ est le nombre de coefficients de l'échelle j ($N^j = \frac{N}{2^j}$ pour $j = 1 \dots M$ et $N^j = \frac{N}{2^M}$ pour $j = M + 1$);
2. calculer le seuil $S_{l+1}^j = f(\sigma_l^j) = F_a \sigma_l^j$, où F_a est un paramètre utilisateur indépendant de l'échelle j ;
3. calculer $\Omega_{r,l+1}^j$ et $\Omega_{\Delta c,l+1}^j$ par seuillage de $\Omega_{r,l}^j$ avec le seuil S_{l+1}^j ;
4. reconstruire \mathbf{r}_{l+1} et $\Delta \mathbf{c}_{l+1}$ à partir des vecteurs $\Omega_{r,l+1}$ et $\Omega_{\Delta c,l+1}$, et fixer $\mathbf{c}_{l+1} = \mathbf{c}_l + \Delta \mathbf{c}_{l+1}$;
5. revenir au pas 1 si le critère d'arrêt (27) n'est pas atteint.

Cet algorithme itératif s'interrompt quand la différence d'énergie sur le bruit résiduel \mathbf{r} reconstruit entre deux itérations successives est jugé suffisamment faible. Un STC_l petit est obtenu quand très peu d'information (parties du signal corrélées avec l'ondelette choisie) est extraite du bruit et ajoutée au signal entre deux itérations successives.

2.4 Interprétation de point-fixe

Nous présentons dans cette section une analyse détaillée de l'algorithme WTST-NST de Hadjileontiadis *et al.* . En effet, l'utilisation d'une base de décomposition-reconstruction unique permet des développements intéressants de cet algorithme. Nous montrons notamment que, dans ce cas, les étapes de décomposition-reconstruction successives (voir figure 3) sont inutiles, ce qui implique une réduction importante de la charge de calcul. L'étude de la convergence de la méthode itérative de seuillage (selon l'équation (28)) nous permet d'éliminer un des paramètres utilisateur ε grâce à l'introduction d'une approche de point-fixe [33] et, de ce fait, d'optimiser l'algorithme de débruitage.

2.4.1 Préliminaires

Les transformées en ondelettes directe (WD) et inverse (WR) appliquées pour les opérations de décomposition et reconstruction sont orthogonales et linéaires. En prenant en compte l'orthogonalité de la transformation en ondelettes, le critère d'arrêt (27) s'écrit :

$$STC_{l+1} = \|\mathbf{r}_l\|^2 - \|\mathbf{r}_{l+1}\|^2 = \|\Omega_{r,l}\|^2 - \|\Omega_{r,l+1}\|^2. \quad (29)$$

En posant $\varepsilon = 0$, les itérations s'arrêtent quand $\Omega_{r,l} = \Omega_{r,l+1}$, c'est-à-dire quand aucun coefficient n'est éliminé par seuillage et $\Omega_{\Delta c,l+1} = \mathbf{0}$. Le critère d'arrêt (29) modifié ($STC_{l+1} = 0$) peut être utilisé à l'étape 5 de l'algorithme précédent, donc la reconstruction du bruit $\mathbf{r}_l = \text{WR}(\Omega_{r,l})$ à chaque itération est inutile.

Grâce à la linéarité de la transformée en ondelettes, le vecteur des coefficients du signal \mathbf{x} s'écrit :

$$\Omega_{\mathbf{x}} = \Omega_{r,l} + \sum_{i=0}^l \Omega_{\Delta c,i} = \Omega_{r,l} + \Omega_{c,l}, \quad \forall l, \quad (30)$$

ce qui implique qu'il n'est pas nécessaire de calculer $\Delta \mathbf{c}_l$ par reconstruction à partir de $\Omega_{\Delta c,l}$ à chaque itération et le signal informatif final \mathbf{c} se reconstruit selon $\mathbf{c} = \sum_{l=0}^L \Delta \mathbf{c}_l = \sum_{l=0}^L \text{WR}(\Omega_{\Delta c,l})$. Ce signal sera calculé seulement après convergence : $\mathbf{c} = \text{WR}(\Omega_{\mathbf{c}}) = \text{WR}(\sum_{l=0}^L \Omega_{\Delta c,l})$, où L est le nombre final d'itérations.

Autrement dit, les étapes intermédiaires de décomposition-reconstruction qui constituent l'étape 4 de l'algorithme WTST-NST peuvent être supprimées et remplacées par une reconstruction finale unique.

¹L'expression du seuil est similaire dans les travaux de Starck *et al.* [30] en imagerie astronomique. Cette valeur de F_a est retrouvée aussi dans la détection des valeurs aberrantes (*outliers*) (voir aussi notre article [32]).

2.4.2 Seuillage

Pour chaque échelle j , le seuillage des coefficients d'ondelette de l'estimation du bruit $\mathbf{\Omega}_{r,l-1}^j$ est effectué avec le seuil $S_l^j = F_a \sigma_{l-1}^j$, qui sépare $\mathbf{\Omega}_{r,l-1}^j$ en $\mathbf{\Omega}_{r,l}^j$ et $\mathbf{\Omega}_{\Delta c,l}^j$. S'il existe au moins un indice q tel que $\mathbf{\Omega}_{\Delta c,l}^j(q) \neq 0$, alors on a $|\mathbf{\Omega}_{\Delta c,l}^j(q)| \geq S_l^j$. En plus, pour tout q , $S_l^j > |\mathbf{\Omega}_{r,l}^j(q)|$ et

$$|\mathbf{\Omega}_{\Delta c,l}^j(q)| \geq S_l^j > |\mathbf{\Omega}_{\Delta c,l+1}^j(q)| \geq S_{l+1}^j > \dots, \quad (31)$$

car $\mathbf{\Omega}_{\Delta c,l+1}^j$ est obtenu par le seuillage de $\mathbf{\Omega}_{r,l}^j$. Par conséquent, tous les coefficients non-nuls de $\mathbf{\Omega}_{c,l}^j = \sum_{i=0}^l \mathbf{\Omega}_{\Delta c,i}^j$ ont un module supérieur ou égal à S_l^j , qui est supérieur ou égal aux coefficients de $\mathbf{\Omega}_{r,l}^j$. En initialisant la suite de seuils par $S_0^j = \infty$ (ce qui est équivalent à l'initialisation $\mathbf{r}_0 = \mathbf{x}$ proposée précédemment), on obtient (voir (31)) une séquence décroissante $S_0^j > S_1^j > \dots > S_{L-1}^j = S_L^j$, où S_L^j est le seuil final obtenu quand $\mathbf{\Omega}_{\Delta c,L}(q) = 0$ quel que soit q (figure 5).

Par conséquent, S_l^j sépare $\mathbf{\Omega}_x^j = \mathbf{\Omega}_{r,l}^j + \mathbf{\Omega}_{c,l}^j$ en deux vecteurs disjoints quel que soit l et donc le seuillage de $\mathbf{\Omega}_{r,l}^j$ par S_{l+1}^j est équivalent au seuillage direct de $\mathbf{\Omega}_x^j$. Les seuillages successifs de $\mathbf{\Omega}_x^j$ sont inutiles et donc le seuil final S_L^j peut être calculé dans une étape antérieure de l'algorithme.

2.4.3 Interprétation de point fixe

Quel que soit le paramètre de seuillage F_a^j , le seuil S_{l+1}^j (à l'itération $l+1$) est calculé en fonction de l'écart-type des coefficients restants dans le bruit à l'itération précédente, c'est-à-dire des coefficients qui respectent $|w_{r,l}^{j,p}| < S_l^j$. En choisissant dans l'éq. (29) $\varepsilon = 0$, l'algorithme s'écrit, pour chaque échelle j :

1. calculer $(\sigma_l^j)^2 = \frac{1}{N_j} \sum_p \left(w_z^{j,p} \mathbb{I}_{(|w_z^{j,p}| < S_l^j)} \right)^2$, où N_j est le nombre de coefficients à l'échelle j et $\mathbb{I}_{(|x| < S)}$ est la fonction indicatrice ($\mathbb{I}_{(|x| < S)} = 1 \quad \forall |x| < S$ et 0 ailleurs) ;
2. calculer S_{l+1}^j comme $S_{l+1}^j = F_a^j \sigma_l^j$;
3. revenir au pas 1. si la convergence n'est pas atteinte.

On peut donc écrire l'itération suivante :

$$S_{l+1}^j = f_j(S_l^j) = F_a^j \sqrt{\frac{1}{N_j} \sum_p \left(w_x^{j,p} \mathbb{I}_{(|w_x^{j,p}| < S_l^j)} \right)^2}, \quad \text{avec } f_j(u) = F_a^j \sqrt{\frac{1}{N_j} \sum_p \left(w_x^{j,p} \mathbb{I}_{(|w_x^{j,p}| < u)} \right)^2}. \quad (32)$$

Évidemment, les itérations $S_{l+1}^j = f_j(S_l^j)$ convergent quand $S_{l+1}^j = S_l^j$, ce qui est équivalent à $S_l^j = f_j(S_l^j)$, donc vers un point fixe de la fonction f_j (voir figure 5(b)).

En conséquence, on opère un débruitage par descente de point fixe pour chaque échelle indépendamment. On interrompt les itérations quand le critère est atteint pour l'échelle donnée. La reconstruction du bruit résiduel et du signal informatif n'est plus nécessaire à chaque itération.

2.4.4 Conditions générales d'application

On a montré au paragraphe précédent que le critère d'arrêt est validé pour un point fixe de la fonction f_j définie antérieurement (32), et que ce point est le seuil final pour l'échelle considérée j . Il reste à étudier dans quelles conditions ce seuil existe, c'est-à-dire les conditions de convergence de l'algorithme ou encore l'existence de ce point fixe de la fonction f_j .

Il est évident que la fonction f_j définie en (32) est une fonction monotone croissante constante par morceaux (voir figure 5(a)), définie pour tout $u \in \mathbb{R}^+$ et à valeurs dans un ensemble fini de nombres réels, étant donné que la dimension du vecteur de coefficients d'ondelette $\mathbf{\Omega}_x^j$ est finie. Les valeurs de la fonction f_j dépendent de la constante utilisateur F_a^j , qui doit respecter certaines conditions pour assurer l'existence du point fixe $f_j(u_f) = u_f$. Selon l'équation (31), l'itération $S_{l+1}^j = f_j(S_l^j)$ génère une suite décroissante de seuils. On peut montrer qu'elle converge vers le premier point fixe rencontré, c'est-à-dire le plus grand :

Notons $\mathbf{\Gamma}$ l'ensemble fini des valeurs de la fonction f_j (32) pour une échelle j donnée, et S_l^j l'ensemble des seuils obtenus par itération ($l \neq 0$). Puisque la suite des seuils est calculée à l'aide de la fonction f_j , $S_l^j \in \mathbf{\Gamma}, \forall l \neq 0$. Supposons que le plus grand point fixe de f_j ne soit pas le point de convergence de l'algorithme. Ceci implique qu'il existe $\gamma = f_j(\gamma)$ tel que $\gamma \in \mathbf{\Gamma}$ et $\gamma \neq S_l^j, \forall l$ (donc un point fixe non-décelé par l'algorithme itératif). Ce constat peut être réécrit : il existe l tel que $S_{l+1}^j < \gamma < S_l^j$.

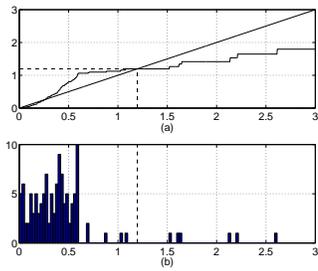


FIG. 5 – Pour une distribution empirique de $w_z^{j,p}$ et un j donné : (a) les fonctions $y = f_j(x)$ (pour $F_a^j = 3$) et $y = x$, montrant l'interprétation de point fixe et la valeur finale $x_0 = S_L^j$ donnée par l'algorithme; (b) l'histogramme des valeurs absolues des coefficients d'ondelette $w_z^{j,p}$. La ligne en pointillé représente le seuil de séparation entre le bruit (les petites valeurs) et le signal informatif.

Mais $S_{l+1}^j = f_j(S_l^j)$, donc la première inégalité s'écrit $f(S_l^j) < \gamma$. D'autre part, f_j est monotone croissante, donc la deuxième inégalité implique $f_j(\gamma) \leq f_j(S_l^j)$, qui implique (car γ est un point fixe) $\gamma \leq f_j(S_l^j)$, ce qui contredit la première inégalité. Par conséquent, l'hypothèse faite au début de notre démonstration est fautive : le plus grand point fixe est le point de convergence de l'algorithme.

L'interprétation de l'algorithme de débruitage comme une descente de point fixe (pour $\varepsilon = 0$) permet de mettre en évidence le rôle du deuxième paramètre utilisateur, la constante multiplicative F_a^j :

- elle doit être choisie suffisamment grande pour assurer la convergence de l'algorithme vers une valeur strictement positive (un seuil nul revient à considérer qu'il n'y a pas de bruit dans le signal \mathbf{x});
- d'autre part, une valeur trop grande de F_a^j conduit à un arrêt prématuré de l'algorithme : si $F_a^j > \frac{\max(|\Omega_x^j|)}{\sigma_0^j}$, alors l'algorithme converge après la première itération (ce qui signifie qu'il n'y a pas de signal informatif dans la bande fréquentielle correspondant à l'échelle courante j).

2.5 Modélisation probabiliste

Dans cette section, nous proposons une méthode de calcul du deuxième paramètre utilisateur F_a^j , ce qui nous permettra de développer une version non-supervisée de l'algorithme de débruitage itératif.

On observe que, dans la plupart des applications réelles, les coefficients d'ondelette sont distribués selon une loi de probabilité symétrique et qui possède un mode en 0. Notons cette densité $p(w^{j,p})$, ou, puisque le développement suivant est identique pour toutes les échelles et pour simplifier la notation, $p(w)$.

Proposition 1. Si la densité de probabilité des coefficients $p(w)$ de moyenne nulle et de variance finie possède un mode en 0 et s'il y a au moins un w réel tel que $p(w) = h(w)$, avec $h(w) = \frac{K}{w}$ et $K = \frac{3}{2F_a^2}$, alors la fonction $f(S) = F_a \cdot \sigma_{|w| < S}$ possède au moins un point fixe (voir figure 6).²

2.5.1 Modèles gaussiens généralisés

Comme nous l'avons montré précédemment, seule une valeur de F_a supérieure à F_{am} assure la convergence de l'algorithme de débruitage pour l'échelle j considérée. Cette valeur minimale dépend de la densité de probabilité des coefficients d'ondelette $p(w)$ ³. Nous nous proposons [32] de déterminer l'expression de F_{am} dans le cas des lois de probabilité gaussiennes généralisées, car ce sont des modèles pertinents pour des données expérimentales. Les densités de probabilité de type gaussienne généralisée sont définies selon :

$$p_{\sigma,u}(w) = \alpha e^{-|\beta w|^u}, \quad (33)$$

$$\text{avec } \beta = \frac{1}{\sigma} \sqrt{\frac{\Gamma(3/u)}{\Gamma(1/u)}}, \quad \alpha = \frac{\beta u}{2\Gamma(1/u)}, \quad \Gamma(u) = \int_0^\infty e^{-x} x^{u-1} dx,$$

où σ désigne l'écart-type, $u > 0$ le paramètre qui gouverne la décroissance de la loi et sa forme ($u = 2$ pour la loi normale et $u = 1$ pour la loi de Laplace).

²Pour la démonstration, voir [33]. La condition est suffisante mais pas nécessaire.

³Pour alléger les notations, on ne fait pas apparaître l'indice j de l'échelle dans le développement présenté ici.

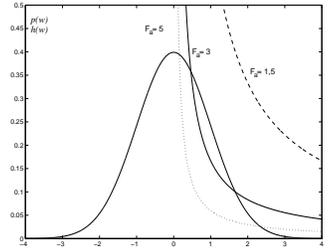


FIG. 6 – Condition d'existence du point fixe. Dans cet exemple, la loi de probabilité $p(w)$ est une gaussienne de moyenne nulle et d'écart-type $\sigma = 1$. La fonction $h(w) = \frac{3}{2F_a^2 w}$ est tracée pour trois valeurs de F_a . La convergence de l'algorithme de point fixe n'est pas assurée pour $F_a = 1, 5$.

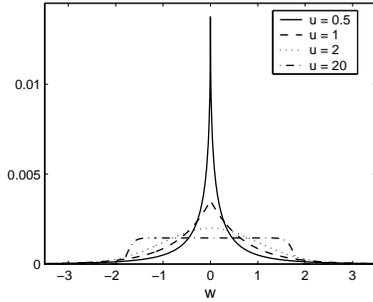


FIG. 7 – Exemples de fonctions de densité de probabilité gaussiennes généralisées $p_{\sigma, u}(w)$ de moyenne nulle et d'écart-type $\sigma = 1$, pour différentes valeurs de u .

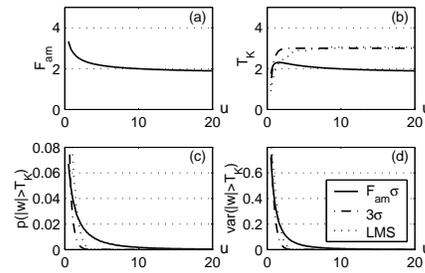


FIG. 8 – Pour différentes valeurs du paramètre u et $\sigma = 1$: (a) évolution de F_{am} ; (b) évolution du seuil final T_K ; (c) probabilité $p(|w| > T_K)$; (d) variance des coefficients w , $|w| > T_K$. Les figures (b)–(d) présentent aussi le seuil, la probabilité et la variance obtenus avec deux autres méthodes de détection de points aberrants : seuillage itératif 3σ et LMS.

On cherche la valeur minimale du paramètre F_a pour laquelle la condition énoncée dans la proposition 1 est respectée. Cela revient à trouver la solution de l'équation $p(w) = h(w)$. L'intersection existe seulement pour w positif et le paramètre β de la gaussienne généralisée est toujours positif, donc l'équation s'écrit :

$$\alpha e^{-(\beta w)^u} = \frac{3}{2F_a^2 w}, \text{ ou encore } w = \frac{3}{2\alpha F_a^2} e^{(\beta w)^u}. \quad (34)$$

La fonction $q(w) = \frac{3}{2F_a^2} e^{(\beta w)^u}$ est monotone croissante et convexe. On montre aisément qu'elle a 0, 1 ou 2 points d'intersection avec la droite $y(w) = w$, et donc que l'équation (34) admet 0, 1 ou 2 solutions. On obtient une solution unique w_0 lorsque $q(w_0)$ est tangente à la première bissectrice. Cela revient à chercher le point w_0 pour lequel on a $q'(w_0) = 1$, c'est-à-dire que :

$$\frac{3}{2\alpha F_a^2} u \beta^u w_0^{u-1} e^{(\beta w_0)^u} = 1. \quad (35)$$

En remplaçant (34) dans (35), on obtient : $w_0^u = \frac{1}{u\beta^u}$, donc $w_0 = \frac{1}{\beta u^{\frac{1}{u}}}$. En remplaçant w_0 en (34), on trouve la valeur de F_{am} minimale pour laquelle cette solution existe :

$$F_{am} = \sqrt{\frac{3\beta}{2\alpha} (ue)^{\frac{1}{u}}} = \sqrt{\frac{3\Gamma(\frac{1}{u})}{u} (ue)^{\frac{1}{u}}}, \quad (36)$$

indépendant de σ mais différent pour chaque échelle j de la décomposition.

2.5.2 Détection de valeurs aberrantes

Le débruitage par ondelettes revient à sélectionner les grandes valeurs (pour reconstruire le signal informatif). En utilisant le seuil calculé à travers la constante F_{am} , on est assuré d'extraire quasiment tous les grands coefficients ou, du point de vue du bruit, de faire une estimation robuste de son écart-type (tous les points aberrants étant attribués au signal). Nous avons donc testé cette méthode en la comparant avec deux méthodes classiques de détection de points aberrants : détection itérative avec une constante $F_a = 3$ et LMS (voir Rousseeuw et Leroy [34]). Les résultats sont détaillés dans [32] (voir aussi figure 8).

2.5.3 Optimisation

Les valeurs minimales F_{am} des constantes multiplicatives assurent la convergence de l'algorithme de débruitage de point fixe vers des seuils différents de 0. Cependant, ces seuils ont un désavantage important : les valeurs de F_{am} sont souvent inférieures à 3 et elles diminuent avec le paramètre u qui gère la forme des gaussiennes généralisées (figures 7 et 8). Dans ce cas, le seuil obtenu est inférieur à celui de Hadjileontiadis *et al.* et la version reconstruite du signal risque de contenir trop de fausses alarmes et donc être trop bruitée. D'autre part, un avantage apporté par les seuils calculés à partir de F_{am} est qu'ils ne dépendent pas de la taille N du signal, comme pour le seuil universel $S_u = \sigma \sqrt{2 \log_e N}$ de Donoho et Johnstone.

Nous avons donc développé trois versions de notre algorithme de point fixe :

1. **seuillage minimal** : $S_l^j = F_{am}^j \sigma_l^j$, avec F_{am}^j valeur minimale du paramètre F_a^j (éq. (36)) qui assure la convergence de l'algorithme de point fixe et σ_l^j écart-type des coefficients d'ondelette du bruit résiduel à l'échelle j et à l'itération l . Ce seuillage conduit en principe à un débruitage sans perte d'information (aucune non-détection, mais certainement de nombreuses fausses alarmes) ;
2. **seuillage "optimal"** : on calcule le paramètre F_{ao}^j tel que les seuils S^j soient initialisés, pour notre algorithme itératif et pour un bruit additif gaussien, aux valeurs correspondant au seuil universel de Donoho et Johnstone :

$$F_{ao}^j = K_o^j F_{am}^j, \quad (37)$$

avec

$$K_o^j = \max(1, K_c^j) \quad \text{avec} \quad K_c^j = 2 \sqrt{\frac{\alpha \log_e N}{3\beta\sqrt{2e}}} = \sqrt{\frac{2u \log_e N}{3\Gamma(\frac{1}{u})\sqrt{2e}}},$$

où N est la dimension du signal.

On vérifie facilement que $F_{ao}^j = \sqrt{2 \log_e N}$ pour les distributions gaussiennes ($u = 2$). L'expression de K_o^j se justifie par le fait que la valeur de K_c^j est inférieure à 1 pour certaines valeurs de u , et donc le paramètre F_a^j calculé à partir de K_c^j devient inférieur à F_{am}^j et la convergence de l'algorithme n'est pas assurée. Comme le seuil universel, l'optimal dépend de la taille N du signal.

3. **seuillage par hystérésis** : ce dernier type de seuillage est une combinaison séquentielle des deux méthodes précédentes. L'algorithme suit trois étapes :
 - (a) on calcule les deux seuils précédents ("optimal" et minimal). Le seuillage "optimal" détecte les grands coefficients d'ondelette, ce qui minimise le nombre de fausses alarmes, alors que le seuillage minimal minimise la perte d'information ;
 - (b) à partir de chaque coefficient retenu par le seuillage "optimal" (qui correspond aux événements importants) on calcule les instants de début et de fin de l'événement détecté ;
 - (c) pour chaque intervalle temporel obtenu, on sélectionne *tous* les coefficients d'ondelette supérieurs au seuil minimal correspondants (dont les ondelettes associées contribuent à la reconstruction de l'événement inclus dans l'intervalle).

Autrement dit, nous utilisons le seuillage "optimal" pour détecter et segmenter rigoureusement les événements importants et minimiser les fausses alarmes, et le seuillage minimal pour garder un maximum d'information possible sur ces événements et assurer leur reconstruction précise.

2.6 Perspectives

En ce moment, nous voyons deux directions de recherche possibles :

- l'approfondissement du seuillage dit par "hystérésis" : le seuil minimal semble remplir sa mission et il a une justification théorique correcte. En revanche, le seuil maximal est obtenu d'une manière *ad hoc*, pour mimer le seuil universel (éq. 37). Nous pensons que ce seuil pourrait être obtenu selon une méthode similaire à celle utilisée pour le seuil universel, c'est-à-dire en faisant appel à la théorie des extrêmes : pour un signal de longueur N , quelle est la valeur du seuil auquel tous les N coefficients sont inférieurs (avec une probabilité qui tend vers 1 quand N grand). Autrement dit, peut-on trouver la loi de la variable $M_j = \max_p(w_x^{j,p})$, et est-ce que cette variable a une espérance finie et une variance, dans le meilleur des cas, tendant vers 0 quand N grand ?
- peut-on utiliser cette méthode de détection de points aberrants dans des cas multi-dimensionaux ? Quelle est la pertinence de l'estimation des paramètres d'une gaussienne généralisée en plusieurs dimensions ? Est-ce que la méthode d'estimation a une influence importante sur le seuil calculé ?

Références

- [1] J. Arnhold, P. Grassberger, K. Lenherz, and C. Elger, "A robust method for detecting interdependencies : application to intracranially recorded eeg," *Physica D*, vol. 134, pp. 419–430, 1999.
- [2] F. Wendling, F. Bartolomei, J. Bellanger, and P. Chauvel, "Interpretation of interdependencies in epileptic signals using a macroscopic physiological model of the eeg," *Clin. Neurophysiol.*, vol. 112, pp. 1201–1218, 2001.
- [3] F. Bartolomei, F. Wendling, J. Bellanger, J. Régis, and P. Chauvel, "Neural networks involving the medial temporal structures in temporal lobe epilepsy," *Clin. Neurophysiol.*, vol. 112, pp. 1746–1760, 2001.

- [4] K. Lehnertz and C. Elger, "Spatio-temporal dynamics of the primary epileptogenic area in temporal lobe epilepsy characterized by neuronal complexity loss," *Electroenceph. Clin. Neurophysiol.*, vol. 95, pp. 108–117, 2000.
- [5] R. Kus, M. Kaminski, and K. Blinowska, "Determination of eeg activity propagation : Pair-wise versus multichannel estimate," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 9, pp. 1501–1510, 2004.
- [6] J. Pardey, S. Roberts, and L. Tarassenko, "A review of parametric modelling techniques for eeg analysis," *Med. Eng. Phys.*, vol. 18, pp. 2–11, 1996.
- [7] C. Bernasconi and P. König, "On the directionality of cortical interactions studied by structural analysis of electrophysiological recordings," *Biol. Cybern.*, vol. 81, pp. 199–210, 1999.
- [8] C. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, pp. 424–438, 1969.
- [9] M. Chavez, J. Martinerie, and M. Le Van Quyen, "Statistical assessment of nonlinear causality : application to epilepticeeg signals," *Journal of Neuroscience Methods*, vol. 124, pp. 113–128, 2003.
- [10] M. Kaminski and K. Blinowska, "A new method of the description of the information flow in the brain structures," *Biol. Cybern.*, vol. 65, no. 3, pp. 203–210, 1991.
- [11] P. Franaszczuk and G. Bergey, "Application of the directed transfer function method to mesial and lateral onset temporal lobe seizures," *Brain Topography*, vol. 11, no. 1, pp. 13–21, 1998.
- [12] —, "An autoregressive method for the measurement of synchronization of interictal and ictal eeg signals," *Biol. Cybern.*, vol. 81, no. 1, pp. 3–9, 1999.
- [13] G. K. Bergey and P. J. Franaszczuk, "Epileptic seizure are characterized by changing signal complexity," *Clin. Neurophysiol.*, vol. 112, no. 2, pp. 241–249, 2001.
- [14] C. Anderson, E. Stolz, and S. Shamsunder, "Multivariate autoregressive models for classification of spontaneous electroencephalogram during mental tasks," *IEEE TBME*, vol. 45, no. 3, pp. 277–286, 1998.
- [15] P. Franaszczuk, K. Blinowska, and Kowalczyk, "The application of parametric multichannel spectral estimates in the study of electrical brain activity," *Biol. Cybern.*, vol. 51, pp. 239–247, 1985.
- [16] H. Akaike, "A new look at statistical model identification," *IEEE trans. automat. control.*, vol. AC-19, pp. 716–723, 1974.
- [17] L. Iasemidis, J. Sackellares, and R. Savit, "Quantification of hidden time dependencies in the eeg within the framework of nonlinear dynamics," in *Nonlinear dynamical analysis of the EEG*, B. Jansen and M. Brandt, Eds. World Scientific : Singapore, 1993, pp. 323–346.
- [18] A. Subasi, M. Kiyimik, A. Alkan, and E. Koklukaya, "Neural network classification of eeg signals by using ar with mle preprocessing for epileptic seizure detection," *Mathematical and Computational Applications*, vol. 10, no. 1, pp. 57–70, 2005.
- [19] S. Marple, *Digital Spectral Analysis with Applications*. Prentice Hall Inc., Englewood Cliffs, New Jersey, 1987.
- [20] M. Takigawa, G. Wang, H. Kawasaki, and H. Fukuzako, "Eeg analysis of epilepsy by directed coherence method. a data processing approach," *Int. J. Psychophysiol.*, vol. 21, pp. 65–73, 1996.
- [21] F. Lopes da Silva, J. Pijn, and P. Boeijinga, "Interdependence of eeg signals : linear vs. nonlinear associations and the significance of time delays and phase shifts," *BrainTopogr.*, no. 2, pp. 9–18, 1989.
- [22] E. Pereda, R. Quiñero, and J. Bhattacharya, "Nonlinear multivariate analysis of neurophysiological signals," *Progress in Neurobiology*, no. 77, pp. 1–37, 2005.
- [23] M. Caparos, V. Louis-Dorr, F. Wendling, J. Vignal, and D. Wolf, "Automatic lateralization of tle based on non-linear correlation on scalp eeg : Study on 43 patients," in *The 3rd European Medical and Biological Engineering Conference. EMBECE'05*, November 2005.
- [24] S. Mallat, *A wavelet tour of signal processing*. Academic Press, 1999.
- [25] Antoniadis, A., Bigot, J., and Sapatinas, T., "Wavelet estimators in nonparametric regression : a comparative simulation study," *Journal of Statistical Software*, vol. 6, no. 6, pp. 1–83, 2001, <http://www-lmc.imag.fr/lmc-sms/Anestis.Antoniadis/HTTP/publis-anto.html>.
- [26] Donoho, D. and Johnstone, I., "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.
- [27] R. Coifman and M. Wickerhauser, "Adapted waveform de-noising for medical signals and images," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, no. 5, pp. 578–586, 1995.

- [28] D. Donoho and I. Johnstone, “Ideal spatial adaptation via wavelet shrinkage,” *Biometrika*, vol. 81, pp. 425–455, 1994.
- [29] S. Berman, *Sojourns and Extremes of Stochastic Processes*. Reading, MA : Wadsworth, 1989.
- [30] J.-L. Starck and F. Murtagh, “Image restoration using wavelet transform,” *Astronomy and Astrophysics*, vol. 288, pp. 342–348, 1994.
- [31] L. Hadjileontiadis and S. Panas, “Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter,” *IEEE Trans. Biomedical Engineering*, vol. 44, no. 12, pp. 1269–1281, 1997.
- [32] Ranta, R., Louis-Dorr, V., Heinrich, C., and Wolf, D., “Iterative wavelet-based denoising methods and robust outlier detection,” *IEEE Signal Processing Letters*, vol. 12, no. 8, pp. 557–560, 2005.
- [33] R. Ranta, C. Heinrich, V. Louis-Dorr, and D. Wolf, “Interpretation and improvement of an iterative wavelet-based denoising method,” *IEEE Signal Processing Letters*, vol. 10, no. 8, pp. 239–241, 2003.
- [34] P. Rousseeuw and A. Leroy, *Robust regression & outlier detection*. John Wiley & sons, inc., 1987.

Photodiagnostic tissulaire *in vivo* : modélisation des interactions lumière-tissus en spectroscopie d'autofluorescence et de diffusion élastique pour identifier les propriétés optiques de tissus biologiques

Walter Blondel*, Emilie Péry, Marine Amouroux, Gilberto Diaz et Cédric Thomas

UMR 7039 Centre de Recherche en Automatique de Nancy (CRAN)

Nancy Université, CNRS, INPL, UHP

*walter.blondel@cran.uhp-nancy.fr

Mots-clés : spectroscopie optique, autofluorescence, diffusion multiple, modélisation des interactions lumière - tissus biologiques, identification de coefficients optiques, simulation statistique

Remerciements

Les auteurs remercient pour leur soutien financier La Ligue Régionale de Lutte Contre le Cancer (CD54, CD52), la région Lorraine et le département Ingénierie du CNRS.

Préambule

L'objectif général de notre projet est de concevoir, développer, caractériser et évaluer de nouvelles modalités de diagnostic par méthodes spectroscopiques *in vivo in situ* (biopsie optique) pour la détection, la localisation et l'identification de cellules ou tissus pathologiques caractérisés par des modifications de leurs propriétés photophysiques.

Nous développons en particulier une approche « multi-modalités » en spectroscopie point par point (autofluorescence, diffusion élastique et gap-FRAP)* et en imagerie (endoscopie de fluorescence multi-bandes)*. Le champ applicatif du projet concerne en première intention la détection de lésions tissulaires épithéliales (cancers de la vessie et de la peau principalement) difficilement décelables par l'observation classique en lumière blanche. Nos travaux en cours concernent 3 axes complémentaires :

- le développement instrumental, avec la réalisation de dispositifs possédant des résolutions spatiale, spectrale et temporelle complémentaires : d'une part, un imageur endoscopique d'autofluorescence multi-bandes fournissant une information spectrale de fluorescence en temps réel sous la forme d'une image superposable à l'image en lumière blanche obtenue classiquement et d'autre part, un spectroscopie bimodal permettant l'acquisition *in vivo* de spectres d'autofluorescence et de diffusion élastique en plusieurs points simultanément.
- la modélisation et l'analyse de données : les travaux en imagerie concernent la mise en oeuvre et l'évaluation de méthodes efficaces de recalage d'images (transformation projective) permettant la réalisation d'une image panoramique des zones observées en endoscopie avec superposition spatiale des points suspects obtenus par les informations d'autofluorescence. L'exploitation des spectres de fluorescence et de diffusion élastique issus de la spectroscopie pour obtenir des critères discriminants, nécessite la mise en oeuvre d'outils mathématiques et de traitement de signal adaptés (modélisation du couplage diffusion-fluorescence, identification, classification). Les résultats attendus sont la validation des différents algorithmes de traitement des images et des spectres à partir de données expérimentales acquises sur fantômes, modèles animaux et tumoraux adaptés...
- la mise en oeuvre complémentaire de ces méthodes pour augmenter la sensibilité et la spécificité de détection des lésions. Nous étudions notamment l'implémentation de la technique de gap-FRAP (Recouvrement de Fluorescence Après Photoblanchiment) en imagerie et en spectroscopie de fluorescence, associée à différents modèles biologiques et mathématiques, dans le but ensuite de la mettre en oeuvre *in situ* à l'échelle tissulaire.

* Cf Définitions en Annexes

* Cf Définitions en Annexes

I.1. INTRODUCTION

I.1.1. Cadre des travaux

Les cancers représentent la 2^{ème} cause de mortalité dans les pays industrialisés. 85% d'entre eux prennent naissance dans les couches superficielles des organes (épithélium), dont 50 % au sein des muqueuses des organes creux (voies aériennes, digestives et urinaires). La détection de ces lésions à un stade précoce constitue un enjeu sociétal, économique et humain majeur (augmentation des chances de guérison, utilisation de traitements moins agressifs, réduction des coûts...). Actuellement, celle-ci est effectuée le plus souvent lors d'un examen endoscopique en lumière blanche (fibroscopie ou vidéo endoscopie), qui permet d'examiner visuellement l'aspect macroscopique de la surface de la muqueuse et de repérer les éventuelles zones suspectes, sur lesquelles seront alors effectués des prélèvements biopsiques permettant de déterminer la nature histo-pathologique exacte du tissu. L'endoscopie conventionnelle permet d'identifier les cancers suffisamment avancés pour avoir provoqué une modification de l'aspect macroscopique de la muqueuse. La visualisation des états précancéreux (dysplasies et carcinomes *in situ*) et cancéreux précoces (cancers micro-invasifs) est en revanche plus difficile et les modifications histologiques ne sont décelables qu'à l'examen microscopique d'un échantillon tissulaire prélevé par biopsie. Le meilleur moyen actuel de déceler la présence de ce type de lésion est donc d'effectuer systématiquement de nombreuses biopsies réparties sur l'ensemble de la muqueuse, sachant que cette procédure a une sensibilité médiocre du point de vue clinique.

La spectroscopie tissulaire est une technique de diagnostic en développement qui s'applique aux tissus biologiques et notamment en cancérologie. Elle fait appel aux interactions de la lumière avec les tissus et consiste à faire l'analyse spectrale du rayonnement émis en réponse à une excitation lumineuse appropriée. L'intérêt clinique de l'approche spectroscopique provient de la nature non ionisante de la lumière, du caractère non traumatique (non invasif) de la mesure et de sa grande sensibilité. En permettant de guider les biopsies, la méthode doit conduire à une amélioration importante de la sensibilité, voire de la spécificité du diagnostic *in vivo*. Différentes méthodes peuvent être envisagées. Historiquement la première à être utilisée à des fins diagnostiques, la spectroscopie de fluorescence reste aujourd'hui la méthode la plus employée *in vivo* : elle donne lieu à la majorité des travaux de recherche en photodiagnostic, tant fondamentaux que cliniques. La spectroscopie de diffusion élastique, malgré un nombre limité de travaux, suscite un intérêt croissant. Ne présentant pas de difficulté technique majeure, son exploitation nécessite en revanche le développement de modèles mathématiques de diffusion de la lumière dans les tissus biologiques tenant compte de leur complexité. La spectroscopie Raman, phénomène de diffusion non élastique, présente le grand intérêt de fournir directement une information sur la composition chimique du tissu. Cependant, les sections efficaces de diffusion Raman sont de plusieurs ordres de grandeur inférieures à celles observées en fluorescence, elles-mêmes très inférieures aux sections efficaces de diffusion élastique ; il en résulte que les signaux Raman ont des intensités très faibles par rapport à la diffusion élastique et à la fluorescence endogène des tissus biologiques ce qui complique fortement la mise en œuvre instrumentale et clinique de la méthode. Les travaux envisagés dans le cadre de notre projet reposent sur l'exploitation de la fluorescence et de la diffusion élastique.

I.1.2. Problématique

Les tissus biologiques constituent des milieux de physiologie et de composition biochimique et anatomique complexes ; ces caractéristiques évoluant notamment avec le temps et les pathologies. Compte tenu de la taille et de l'organisation des éléments cellulaires et structuraux des tissus (quelques dizaines de nanomètres à quelques centaines de

micromètres) ainsi que de leur nature chimique, les interactions de la lumière avec ces milieux turbides font intervenir de nombreux processus. Dans le domaine de l'UV (Ultra-Violet) au proche IR (Infra-Rouge), les mécanismes les plus importants sont l'absorption, la réflexion, la diffusion élastique et inélastique et la photoluminescence (fluorescence, phosphorescence). Ils sont la base de nombreuses méthodes de microscopie et de spectroscopie qui visent à caractériser la nature chimique d'un milieu à partir de l'analyse spectrale des rayonnements absorbés ou émis par celui-ci, en réponse à une excitation lumineuse appropriée.

Dans le cas d'échantillons « minces » (quelques micromètres), l'intensité de fluorescence est notamment proportionnelle à la concentration et au rendement quantique de fluorescence. Dans un milieu plus épais, très diffusant et optiquement inhomogène comme un tissu biologique, les processus de diffusion jouent un rôle primordial dans la propagation de la lumière, et les parcours de diffusion des photons dans l'échantillon doivent être pris en compte. Porteuses d'informations liées respectivement à la nature biochimique et à la structure du tissu (tailles et formes des éléments constitutifs, orientation, coefficients optiques...), l'autofluorescence et la diffusion multiple des photons « déforment » de manière significative les caractéristiques d'émission spectrale des tissus biologiques. L'intensité lumineuse mesurée (spectre, image) est donc la résultante d'une combinaison d'absorption « pure », d'absorption pour la fluorescence (excitation) et de diffusion multiple, qui dépendent toutes de la longueur d'onde, parfois de manière non monotone. Par exemple, pour le diagnostic, l'intérêt d'un ou plusieurs fluorophores intrinsèques sera fonction entre autre de leur localisation tissulaire ou cellulaire (profondeur, concentration), de leurs spectres d'excitation – émission voire des matrices spectrales d'excitation – émission, ainsi que de l'organe étudié.

Pour extraire ces informations et interpréter les résultats expérimentaux, il est nécessaire de développer un modèle optique approprié aux tissus biologiques et de résoudre le problème de propagation du rayonnement dans ce milieu. La théorie du transfert radiatif sert de base à la photométrie et des solutions exactes à l'Equation de Transfert Radiatif (ETR) existent pour un petit nombre de cas particuliers (géométrie plan parallèle, diffusion isotrope). Quelques approximations sont classiquement utilisées en optique tissulaire (approximations au 1^{er} ordre, de diffusion, aux petits angles, théorie des flux...). Le développement de nouvelles méthodes pour solutionner les problèmes de transfert radiatif direct et inverse pour des milieux aux configurations et conditions aux limites arbitraires est crucial.

Les méthodes de simulation statistique de type Monte Carlo (MC) apparaissent particulièrement prometteuses pour la simulation numérique de l'ETR en optique tissulaire. La méthode consiste à simuler le transport de photons dans un milieu complexe en imitant leur migration aléatoire dans l'échantillon, suivie depuis leur entrée jusqu'à leur absorption totale ou leur sortie. L'intérêt de la méthode réside dans sa grande souplesse de paramétrage avec notamment : la prise en compte de vecteurs ou de matrices d'absorption et de diffusion spécifiques, l'utilisation de géométries d'illumination particulières et de sites de détection en nombre quelconque, la mise en œuvre de géométries tissulaires « complexes » (multicouches...), la prise en compte de différents phénomènes optiques dont la fluorescence, la simulation en régime d'excitation impulsionnelle ou continue... Bien que la méthode permette de déterminer la distribution 3D de la lumière dans le milieu, la principale limitation concerne le temps de calcul et la précision du résultat final du fait de l'erreur statistique inhérente au processus de simulation.

Chaque nouveau code doit être validé sur la base, par exemple, de jeux de données expérimentales provenant de fantômes mixtes (absorbants, diffuseurs, fluorophores). Plusieurs algorithmes MC ont déjà été validés pour simuler le transport de lumière de fluorescence dans des milieux absorbants et fortement diffusants ; en faisant notamment varier la concentration en fluorophore et/ou les coefficients optiques du milieu, mais en

prenant en compte jusqu'à présent un seul fluorophore (par couche) et un nombre très réduit de longueurs d'ondes particulières.

Après avoir rappelé le problème de la modélisation expérimentale des caractéristiques d'un tissu par des méthodes optiques, les caractéristiques spectrales des interactions lumière – tissu et le principe de la simulation statistique, nous présentons une solution algorithmique pour le couplage absorption – fluorescence multiple dans un milieu multicouches. La solution proposée permet de prendre en compte plusieurs fluorophores par couches ainsi qu'un échantillonnage spectral complet des paramètres optiques dans la bande proche UV – proche IR. Nous présentons finalement quelques résultats de simulation montrant le fonctionnement conforme de l'algorithme, vis-à-vis de l'expérimentation mais aussi mettant en évidence la difficulté, compte tenu des temps de calcul inhérents, d'utiliser l'approche de simulation actuelle directement dans un processus d'identification.

I.2. MATÉRIELS ET MÉTHODES

I.2.1. Configuration expérimentale et modèle générique

La figure ci-après présente le schéma de principe du système expérimental de spectroscopie bimodale mis en oeuvre :

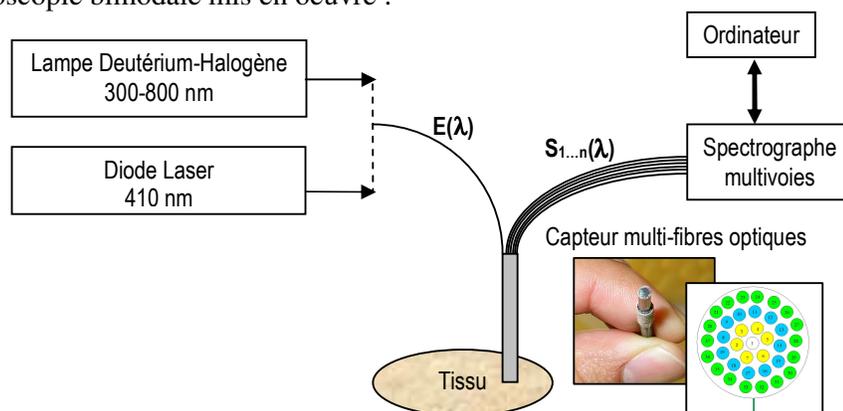


Figure 1. Schéma de principe du système expérimental de spectroscopie bimodale

Deux sources de lumière continues bande étroite (diode laser pour l'excitation de fluorescence) et bande large (lampe Deutérium-Halogène pour la diffusion élastique) permettent alternativement d'illuminer le tissu au travers d'une des fibres optiques du capteur. Ce capteur possède 37 fibres optiques à différentes distances les unes des autres dont 13 sont connectées à un spectrographe imageant permettant l'acquisition simultanée des 13 spectres d'intensité émis par le tissu à ces différentes distances.

Pour modéliser le comportement optique des tissus, un modèle générique stratifié est utilisé (cf. figure 2 ci-après) dans lequel chaque couche homogène est décrite par un ensemble de paramètres caractéristiques (épaisseur, indice de réfraction, coefficients d'absorption, de diffusion et d'anisotropie, spectres d'excitation et d'émission des fluorophores, concentrations moléculaires, etc.).

Le schéma classique d'identification correspondant est donné ci-après ; dans le cas du diagnostic *in vivo*, il doit servir à déterminer les valeurs spectrales d'un ou de plusieurs paramètres cités précédemment dans une ou plusieurs couches.

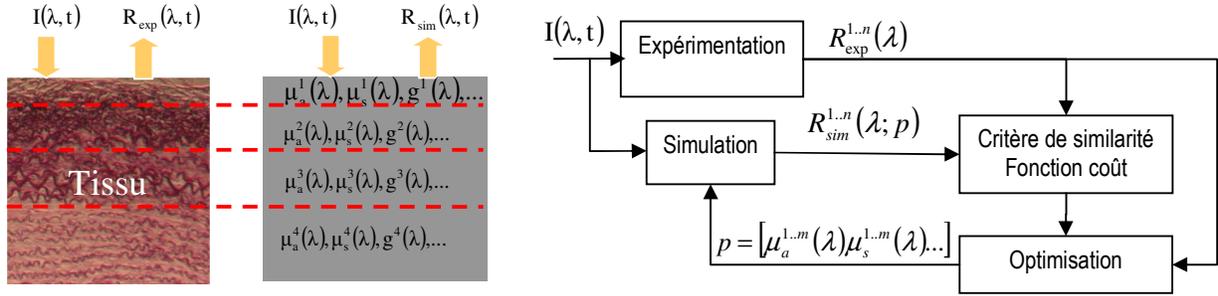


Figure 2. Schémas du modèle optique stratifié de tissu (gauche) et du processus d'identification associé (droite). $I(\lambda, t)$ et $R(\lambda, t)$ sont respectivement les intensités spectrales incidente et rétrodiffusées et $\mu_a^i(\lambda)$, $\mu_s^i(\lambda)$ et $g^i(\lambda)$ les coefficients d'absorption, de diffusion et d'anisotropie de la couche i du modèle de tissu. p est le vecteur des paramètres à optimiser.

I.2.2. Modèles physiques d'absorption et de diffusion

La bande de longueur d'onde d'intérêt pour la spectroscopie d'autofluorescence et de diffusion élastique s'étend de l'UV au proche IR (300 – 1000 nm) et permet donc d'atteindre des profondeurs de pénétration de quelques 10 μm à quelques cm (IR). Dans cette bande, on trouve une fenêtre optique (550 – 1000 nm) d'intérêt pour le photodiagnostic compte tenu de l'absorption limitée de l'eau et de l'hémoglobine. Pour ces gammes de longueurs d'onde, les tissus biologiques sont absorbants, fluorescents et diffusants. Ils constituent des milieux hétérogènes à l'échelle microscopique (fibres structurales, cellules, organites cellulaires...) mais pouvant être considérés comme homogène à l'échelle macroscopique.

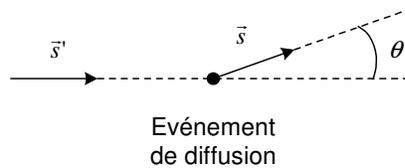
Nous caractérisons quantitativement l'absorption et la diffusion dans un milieu donné en exprimant les paramètres suivants :

- le coefficient d'absorption μ_a (cm^{-1}) qui caractérise l'absorption à l'intérieur d'un milieu non diffusant, homogène, tel que (loi de Beer-Lambert) :

$$I = I_0 \exp(-\mu_a L) \quad (1)$$

où I_0 est l'intensité incidente et I l'intensité transmise.

- La fonction de phase $f(\vec{s}', \vec{s})$ et le facteur d'anisotropie g . La fonction de phase décrit la probabilité qu'un photon de direction incidente \vec{s}' subissant une diffusion, poursuive son déplacement dans une direction \vec{s} formant un angle θ par rapport à \vec{s}' .



Pour les tissus biologiques, la fonction de phase peut être exprimée en fonction du cosinus de l'angle θ de rediffusion $f(\vec{s}', \vec{s}) = f(\cos \theta)$. La fonction la plus utilisée pour modéliser la diffusion dans les tissus est la fonction de Henyey-Greenstein :

$$f(\cos \theta) = \frac{1}{4\pi} \frac{1 - g^2}{(1 + g^2 - 2g \cos \theta)^{3/2}} \quad (2)$$

Le coefficient d'anisotropie g est défini comme le cosinus moyen de l'angle de rediffusion :

$$g = \int_{-1}^{+1} f(\cos\theta) \cos\theta d(\cos\theta) \quad (3)$$

$g = 0$, >0 et <0 correspond à une diffusion isotrope, vers l'avant et vers l'arrière.
 $0.8 < g < 0.98$ pour les tissus biologiques.

- les coefficients de diffusion μ_s et de diffusion réduit μ_s' (cm^{-1}) :

$$\mu_s' = (1 - g)\mu_s = \frac{1}{l^*} = \frac{1 - g}{l_s} \quad (4)$$

où l_s et l^* sont respectivement les libres parcours moyens de diffusion (20-100 μm) et de transport ($\sim 500 \mu\text{m}$).

I.2.3. Théorie du Transfert Radiatif

L'Equation de Transfert Radiatif (ETR) ou équation de transport de densité d'énergie radiative (émission) décrit la propagation d'une énergie lumineuse dans un milieu caractérisé par ses paramètres macroscopiques (μ_a , μ_s , g). Elle décrit le transport particulaire d'un rayonnement dans la direction $\vec{s} \in S^2$ (sphère unitaire dans \mathfrak{R}^3) sous la forme de variations spatiales (au point $r \in \mathfrak{R}^3$ représenté par \vec{r}) et temporelles $t \geq 0$ locales de la luminance $L(\vec{r}, \vec{s}, t, \lambda)$ également fonction de la longueur d'onde $\lambda > 0$.

$$\frac{1}{c} \partial_t L + \vec{s} \cdot \vec{\nabla}_r L = \mu_s \int_{4\pi} f(\vec{s}', \vec{s}) L d\vec{\Omega} - (\mu_a(\lambda) + \mu_s(\lambda)) L + S(\vec{r}, \vec{s}, t, \lambda) \quad (5)$$

Avec $L = L(\vec{r}, \vec{s}, t, \lambda)$. Les termes situés à gauche de cette égalité décrivent respectivement les variations temporelles et spatiales de la luminance dans la direction \vec{s} . Les termes de droites représentent les différentes contributions équilibrant cette équation. Le premier terme correspond à la somme de tous les photons provenant de toutes les directions \vec{s}' et re-diffusés dans la direction \vec{s} . Le deuxième terme (négatif) décrit les pertes de photons dues à l'absorption ($\mu_a(\lambda)$) et à la diffusion ($\mu_s(\lambda)$) en fonction de la longueur d'onde. Enfin, le troisième terme permet de tenir compte de sources locales de photons (par exemple des fluorophores extrinsèques ou intrinsèques).

La forme de cette équation implique d'envisager différents cas de résolutions :

- $\mu_s = 0$, on peut trouver des solutions analytiques (avec $L(\vec{r}, \vec{s}, t, \lambda) = L(\vec{r}, \vec{s}, \lambda)$ continues ou $S(\vec{r}, \vec{s}, t, \lambda) = 0$), mais ce cas ne correspond pas à la situation des tissus biologiques qui sont au contraire très diffusants.
- $\mu_s \neq 0$, avec deux approches possibles :
 - résolution numérique. Par exemple, les méthodes des flux (méthodes multi-flux, méthode P_N , méthodes des ordonnées discrètes) séparent les dépendances angulaires et spatiales de façon à remplacer l'ETR par un système d'équation aux dérivées partielles (EDP) après intégration sur tout l'espace angulaire. Néanmoins, différentes hypothèses sont associées à cette approche où l'on considère notamment une diffusion isotrope et $\mu_a \ll \mu_s$, ce qui limite également son application dans le cas des tissus biologiques.
 - simulation numérique (détaillée ci-après), autrement appelée propagation de rayons ou migration de photons dans les applications des interactions lumière-matière.

I.2.4. Principe de simulation statistique des interactions lumière-tissus

Parmi les différentes méthodes mises en œuvre pour modéliser la propagation de la lumière dans les tissus biologiques, la simulation statistique (MC) est actuellement la méthode la plus étudiée pour obtenir une solution (version) numérique de l'Equation de Transfert Radiatif dans le cas de ces milieux complexes et inhomogènes.

Le schéma de la Figure 3 illustre le principe de migration aléatoire de photons qui consiste, pour un très grand nombre de photons (typiquement au moins quelques dizaines de milliers), à faire subir à chacun un déplacement dans des directions et sur des distances tirées au sort suivant des fonctions de densités de probabilités caractéristiques du milieu (μ_a , μ_s , g). Par ailleurs, afin de réduire le nombre de photons à envoyer, chaque photon est en fait considéré comme un paquet de photons qui perd une partie de son énergie totale à chaque évènement d'absorption jusqu'à un certain seuil au dessous duquel il n'est plus propagé (technique de réduction de la variance).

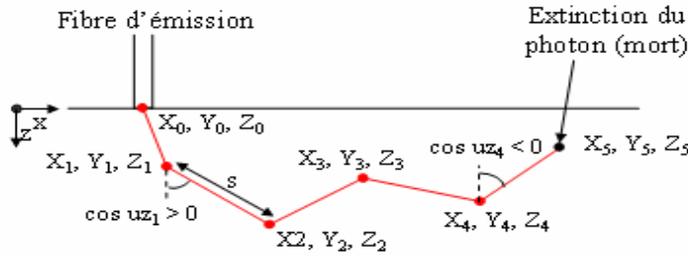


Figure 3. Principe de migration aléatoire de photons

Brièvement, chaque photon est lancé à partir de la position d'injection (X_0, Y_0, Z_0) et possède un « poids » de 1. Chaque nouvelle direction 3D de la trajectoire du photon est déterminée par un échantillonnage statistique (tirage au sort avec w uniformément distribué entre 0 et 1) :

- de l'angle de déflexion θ dans $[0, \pi]$, c'est-à-dire l'angle entre les droites de la précédente et de la nouvelle direction dans le même plan, défini par son cosinus selon :

$$\cos \theta = \frac{1}{2g} \left(1 + g^2 - \left(\frac{1 - g^2}{1 - g + 2gw} \right)^2 \right) \quad (6)$$

- et de l'angle azimutal ϕ dans $[0, 2\pi]$, c'est-à-dire l'angle avec le plan précédent défini simplement par :

$$\phi = 2\pi w \quad (7)$$

La distance parcourue dans la nouvelle direction (pas variable Δs) est également tirée au sort selon une fonction de densité de probabilité dépendant des paramètres optiques de la couche tel que :

$$\Delta s = - \frac{\ln w}{\mu_t} \quad (8)$$

Après chaque déplacement (pas), le photon dépose une partie de son « poids » par absorption aux coordonnées auxquelles il se trouve dans le milieu et proportionnel au rapport des coefficients optiques. La course de chaque photon se termine soit lorsque, ayant subit un nombre suffisant d'absorption, son poids n'est plus suffisant pour lui permettre de poursuivre, soit lorsqu'il sort du milieu. Les poids des photons sortant du milieu aux coordonnées correspondantes aux positions des fibres d'émission sont finalement comptabilisés.

D'autres phénomènes d'interaction lumière-tissu peuvent également être pris en compte et implémentés dans le cadre de ce type d'algorithme tels que : la fluorescence ou la diffusion

inélastique (Raman). Pour notre application (excitation continue bande étroite ou large, faible puissance), l'émission de fluorescence doit être prise en compte en relation avec les caractéristiques spectrales des fluorophores présents dans les différentes couches (différentes profondeurs) du tissu.

1.2.5. Implémentation de la fluorescence multiple

Les principaux travaux déjà réalisés [Liu-2003] ont permis de valider un algorithme de simulation intégrant le couplage en fluorescence pour un modèle monofluorophore et monocouche. A partir des paramètres μ_{af} (coefficient d'absorption du fluorophore) et ϕ (rendement quantique de fluorescence) pour la simulation, les auteurs définissent une probabilité d'absorption pour la fluorescence. Une validation expérimentale est montrée sur fantômes (encre, billes de polystyrène, FAD) à différentes concentrations, avec une excitation à 460 nm et une mesure en émission à 520 nm. Le tableau ci-après récapitule les caractéristiques des principaux programmes existants et celles de l'algorithme développé au laboratoire.

Tableau 1 : Comparaison de différentes simulations (Mcml : [Wang, 1995], Mcmf : [Liu, 2003], notre approche)

Caractéristiques	Références		
	Mcml	Mcmf	Notre approche
Langage de programmation	C	C	C++
Simulateur	Multi-couches	Multi-couches	Multi-couches
Absorption / Diffusion	Propagation d'énergie	Méthode Tout ou Rien	Propagation d'énergie
Prise en charge du spectre complet	-	-	oui
Fluorescence	1 fluorophore	1 fluorophore	Plusieurs fluorophores Effet cascade Résolution temporelle

Dans l'algorithme que nous avons implémenté (cf figure 4), l'excitation des fluorophores est prise en compte en fonction de leurs spectres d'absorption et de leurs rendements quantiques (qui peuvent être différents pour chaque couche). Les paramètres d'absorption et d'émission, le rendement quantique et le coefficient d'absorption du fluorophore sont stockés dans un fichier spécifique pour chaque fluorophore. Le tirage au sort des longueurs d'onde d'émission de fluorescence s'effectue en fonction de la courbe d'émission paramétrée du fluorophore. Nous partons du principe que chaque couche du milieu étant homogène, elle peut contenir une certaine concentration en fluorophore et ainsi, une molécule fluorescente aura plus ou moins de chance d'émettre une fluorescence selon sa concentration relative. Si l'on paramètre plusieurs fluorophores dans une couche, un tirage aléatoire également fonction des concentrations entre les différents fluorophores est réalisé pour déterminer le fluorophore absorbant. En fonction du spectre d'absorption du fluorophore, un coefficient A (cf. équations 9 et 10) est calculé comme probabilité du photon d'être absorbé par le fluorophore à chaque événement :

$$A = \frac{\mu_{af}}{\mu_{af} + \mu_{sf}} \quad (9)$$

$$\mu_{sf} = \mu_s \quad (10)$$

où μ_{af} , μ_{sf} représentent respectivement les coefficients d'absorption et de diffusion du fluorophore considéré et μ_s le coefficient de diffusion du milieu.

Lorsqu'un évènement d'absorption d'un photon propagé se produit quelque part dans le milieu, deux cas sont possibles :

- Si ce photon est absorbé par un chromophore (absorbeur pur) ou par un fluorophore mais sans émission de fluorescence, le niveau d'énergie du photon est mis à jour en conséquence et celui-ci continue son parcours, jusqu'au prochain évènement.
- Si ce photon est absorbé par un fluorophore avec émission de fluorescence, il dépose sur place la totalité de son énergie et donne « naissance » à un photon de fluorescence à l'endroit où il a été absorbé. Ce nouveau photon, dont la longueur d'onde est choisie aléatoirement d'après la courbe d'émission du fluorophore, est alors propagé en considérant une émission isotrope c'est-à-dire dans une nouvelle direction totalement aléatoire (l'angle de ré-émission est tiré au sort entre 0° et 360°).

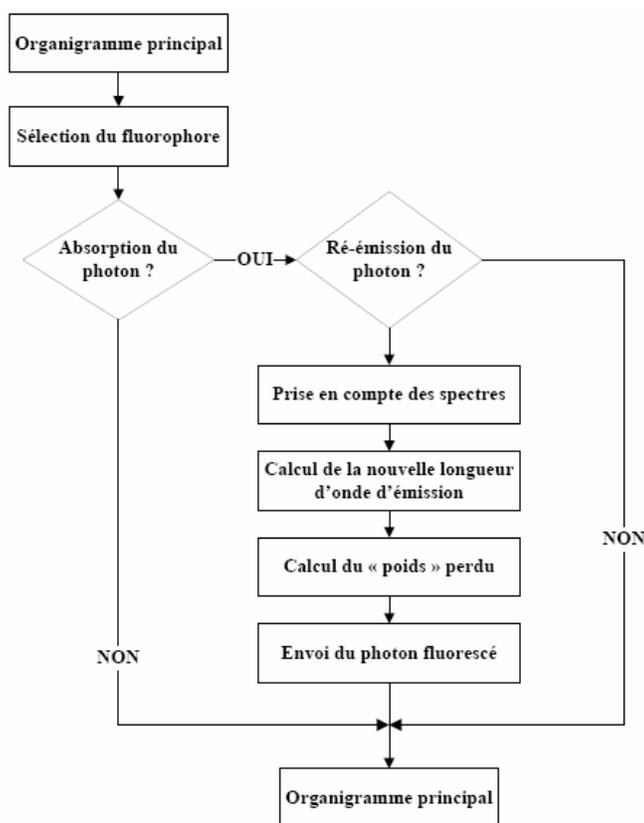


Figure 4. Organigramme du programme de simulation de l'émission de fluorescence multiple

I.3. RESULTATS ET DISCUSSION

Différents tests d'épreuve ont été réalisés avant de travailler sur des fantômes. Nous avons effectué un grand nombre de simulations avec différents profils spectraux de coefficients d'absorption et de diffusion. Pour valider le comportement de l'ensemble du programme, nous avons comparé les spectres simulés et ceux expérimentaux obtenus avec le dispositif expérimental sur de nombreux fantômes mono et multi-couches (absorption, diffusion, fluorescence) : intralipides, encre de chine, fluorophores émettant dans différentes bandes spectrales (fluorescéine, cryptocyanine, éosine).

I.3.1. Exemples de résultats de simulation en absorption / diffusion

- **Simulation n°1**

Excitation : large bande uniforme entre 600 et 800 nm.
Mesures : 2 mesures à 0.53 et 1.74 mm de distances
Nombre de photons / λ : 500 K
Nombre de longueurs d'ondes simulées : 20
Couches : 1 seule de 6 mm

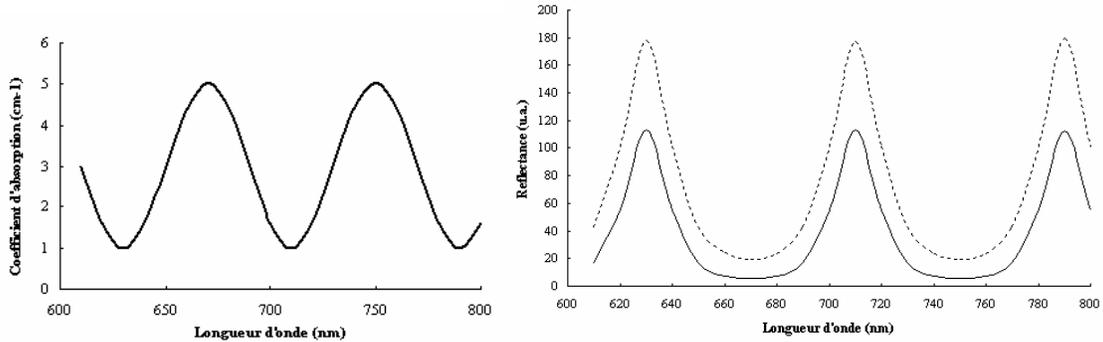


Figure 5. $\mu_s = 100 \text{ cm}^{-1}$, μ_a sinusoïdal (gauche), réflectances simulées (0.53 / 1.74 mm respectivement en traits pointillé / continu)

Durée de simulation : 47 min.

Sur la figure 5, on peut noter les pics et creux d'intensité du spectre de réflectance aux longueurs d'onde correspondant (avec une déformation non-linéaire) respectivement aux creux et pics des valeurs du coefficient d'absorption. On remarque également que l'intensité moyenne (i.e. le nombre de photons) est plus faible pour la distance de fibre la plus grande.

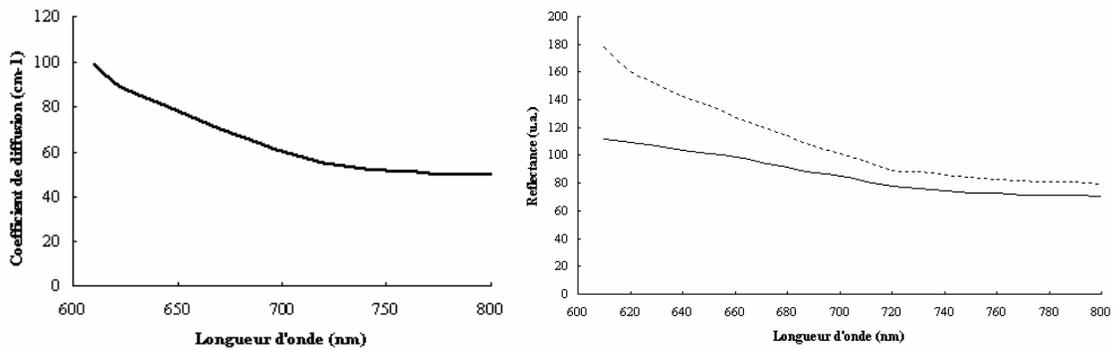


Figure 6. μ_s exponentiel décroissant (gauche), $\mu_a = 1 \text{ cm}^{-1}$, réflectances simulées (0.53 / 1.74 mm respectivement en traits pointillé / continu)

Durée de simulation : 43 min.

La figure 6 montre la diminution de la réflectance en fonction de la longueur d'onde conformément à la diminution de la valeur du coefficient de diffusion en fonction de la longueur d'onde. On peut noter que si l'intensité de réflectance moyenne est plus faible pour la fibre la plus éloignée, l'allure des 2 courbes est légèrement différente avec une diminution plus marquée pour la distance la plus courte. Les photons de longueur d'onde 620 nm se propageant moins loin que ceux de longueur d'onde 780 nm.

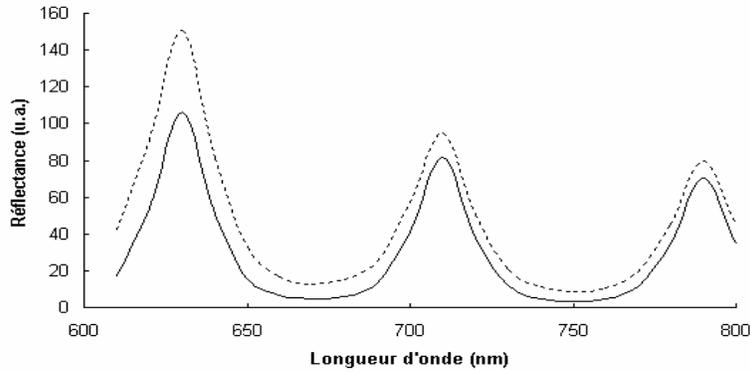


Figure 7. μ_s exponentiel décroissant, μ_a sinusoïdal, réflectances simulées (0.53 / 1.74 mm respectivement en traits pointillé / continu)

Durée de simulation : 30 min.

La figure 7 met en évidence l'ensemble des remarques formulées pour les figures 5 et 6.

• **Simulation n°2**

- Excitation : large bande uniforme entre 600 et 800 nm.
- Mesures : 2 mesures à 0.53 et 1.74 mm de distances
- Nombre de photons / λ : 500 K
- Nombre de longueurs d'ondes simulées : 20
- Couches : 1 couche de 5 mm purement absorbante (μ_a constant)
- 1 couche de 5 mm purement diffusante (μ_s exponentiel décroissant)
- 1 couche de 5 mm purement absorbante (μ_a sinusoïdal)

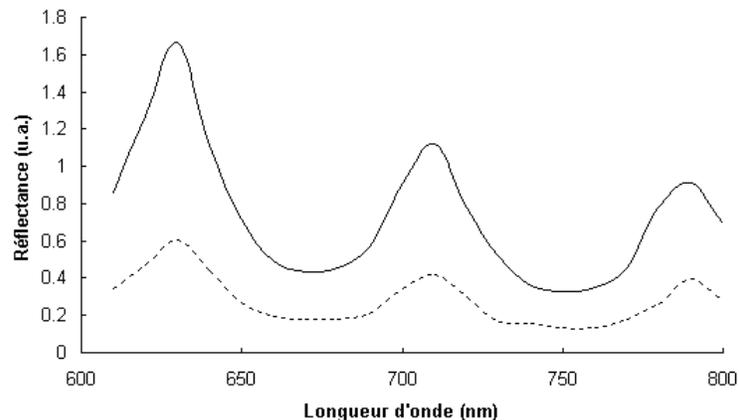


Figure 8. Réflectances simulées (0.53 / 1.74 mm respectivement en traits pointillé / continu) pour un modèle multicouches : 1 couche de 5 mm purement absorbante ($\mu_a = 1 \text{ cm}^{-1}$), 1 couche de 5 mm purement diffusante (μ_s exponentiel décroissant) et 1 couche de 5 mm purement absorbante (μ_a sinusoïdal)

Durée de simulation : 8 min.

La figure 8 illustre de façon extrême (couches purement absorbantes et diffusante) la manière dont une intensité lumineuse incidente peut se propager dans un milieu et être ré-émise avec moins d'intensité à proximité du point d'injection et plus d'intensité à d'autres distances plus grandes. La seconde couche joue ici le rôle équivalent d'une sorte de guide d'onde.

1.3.2. Exemples de résultats de simulation couplant la fluorescence

- **Simulation n°3 : fluorophore unique simulé**

Excitation : bande étroite pic d'excitation à 610 nm.
 Mesures : 2 mesures à 0.53 et 1.74 mm de distances
 Nombre de photons / λ : 500 K
 Nombre de longueurs d'ondes simulées : 20
 Couches : 2 couches de 3 mm chacune

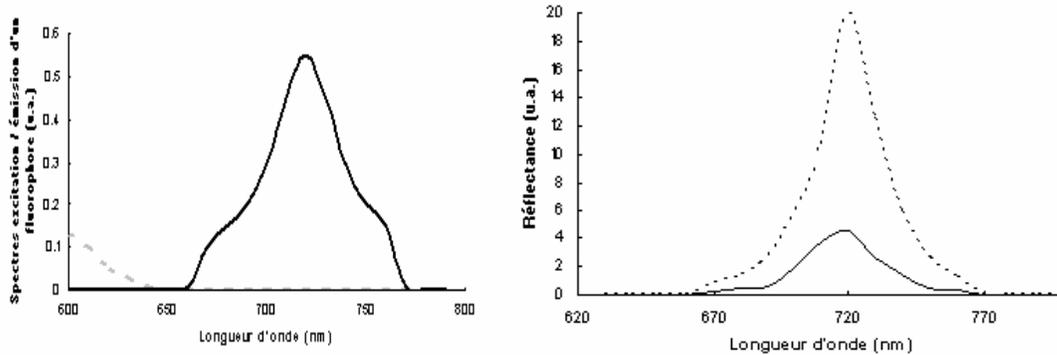


Figure 9. Spectre d'excitation / émission du fluorophore (gauche),
 Réflectances simulées (0.53 / 1.74 mm respectivement en traits pointillé / continu)
 Fluorophore en couche 1, $\mu_a = 1 \text{ cm}^{-1}$ et $\mu_s = 100 \text{ cm}^{-1}$.

La figure 9 montre que la forme du spectre de réflectance ainsi que la valeur de la longueur d'onde du pic d'intensité correspondent aux caractéristiques du spectre d'émission du fluorophore simulé, dans le cas où les coefficients optiques restent constants en fonction de la longueur d'onde.

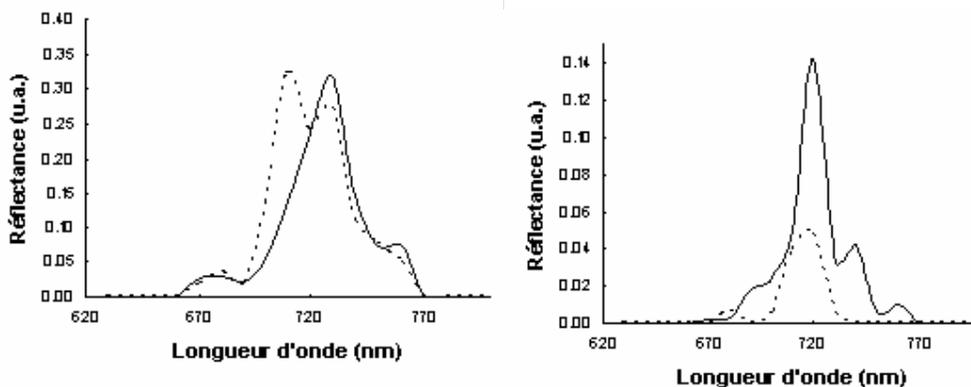


Figure 10. Réflectances simulées (0.53 / 1.74 mm respectivement en traits pointillé / continu)
 Gauche : Fluorophore en couche 1, μ_a sinusoïdal et μ_s exponentiel décroissant
 Droite : Fluorophore en couche 2, $\mu_a = 1 \text{ cm}^{-1}$ et $\mu_s = 100 \text{ cm}^{-1}$

Les courbes de la figure 10 mettent en évidence la manière dont la forme et la valeur du pic en longueur d'onde peuvent être modifiées par les variations spectrales des absorbeurs et diffuseurs associés (déplacement du pic dû aux variations sinusoïdales de μ_a) ou bien par le fait que le fluorophore soit présent seulement dans une couche inférieure (intensité supérieure pour la distance de fibre la plus grande).

• **Simulation n°4 : validation expérimentale avec 1 fluorophore**

Fluorophore : Eosine, éthanol standard
 Concentration : 10 μ Mol
 Excitation : pic d'excitation à 490 nm.
 Mesures : 2 mesures à 0.53 et 1.74 mm de distances
 Nombre de photons / λ : 500 K
 Nombre de longueurs d'ondes simulées : 40
 Couches : 1 couche de 5 mm

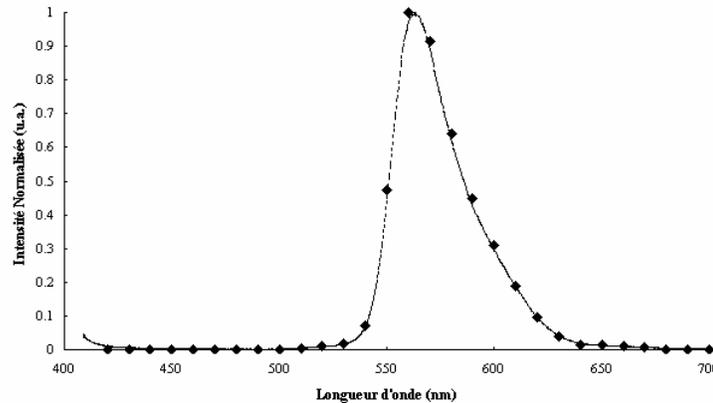


Figure 11. Réflectances mesurée expérimentalement (trait continu) et simulée (points)
 Durée de simulation : 5 min.

La figure 11 montre la bonne corrélation entre le spectre de réflectance obtenu expérimentalement (fantôme liquide) avec le dispositif de mesure et les valeurs théoriques obtenues par simulation en utilisant le spectre d'émission connu du fluorophore (éosine).

• **Simulation n°5 : validation expérimentale avec 2 fluorophores (fantôme liquide)**

Fluorophores : Fluorescéine, éthanol standard & Cryptocyanine, éthanol standard
 Concentration : 10 μ Mol
 Excitation : pic d'excitation à 425 et 650 nm.
 Mesures : 2 mesures à 0.53 et 1.74 mm de distances
 Nombre de photons / λ : 500 K
 Nombre de longueurs d'ondes simulées : 40
 Couches : 1 couche de 5 mm

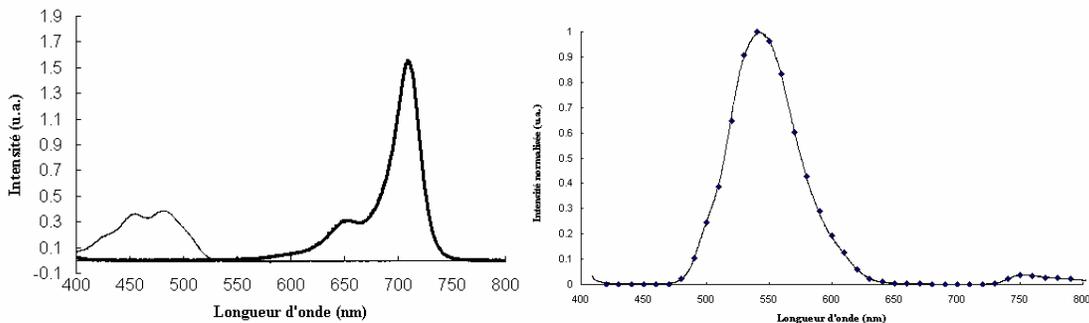


Figure 12. Gauche : Spectres d'absorption de la fluorescéine (trait fin) et de la cryptocyanine (trait épais). Droite : Réflectances mesurée expérimentalement (trait continu) et simulée (points) pour un fantôme liquide composé de 2 fluorophores.

Durée de simulation : 5 min.

L'ensemble des simulations a été effectué en utilisant un pas d'échantillonnage de 10 nm qui convient pour des courbes spectrales variant avec de faibles dérivées. Néanmoins, une diminution de ce pas d'échantillonnage à quelques nm permettrait de mieux représenter les parties les plus abruptes de certaines courbes de réponse (par exemple figure 10).

500 mille photons ont été lancés pour chaque longueur d'onde de simulation, soit un total de 10 millions de photons pour la bande spectrale simulée avec un pas de 10 nm. Un trop grand nombre de photons lancés entraîne un temps de calcul trop élevé pour être acceptable tandis qu'une trop faible quantité de photons lancés entraîne des résultats peu reproductibles (bruit statistique).

Par ailleurs, les durées de simulation obtenues peuvent varier de façon très importante selon les valeurs des coefficients optiques. Ainsi, si la précision sur la solution finale est proportionnelle à $1/\sqrt{N_p}$ avec N_p le nombre de photons propagés (Prah1-1998), le nombre de photons à lancer va, lui, dépendre des propriétés du milieu c'est-à-dire des valeurs des coefficients optiques. En moyenne, pour les tests réalisés, nous avons estimé que 500 000 photons étaient nécessaires par longueur d'onde en fluorescence et 35 000 photons par longueur d'onde en réflectance diffuse pour obtenir un résultat reproductible, nombre de photons qu'il est possible de diminuer ou nécessaire d'augmenter selon les caractéristiques du tissu.

Les spectres d'intensité de réflectance diffuse et d'autofluorescence sont porteurs d'informations sur l'état ou la nature du tissu et l'identification « correcte » des paramètres optiques d'un modèle de ce dernier est donc capitale.

Nos travaux en cours concernent la modélisation expérimentale et en particulier l'identification des coefficients optiques (cf schéma de la figure 2) de tissus sains et tumoraux à partir d'une part, de données expérimentales acquises sur différents modèles *in vivo* et d'autre part, de l'utilisation de l'algorithme de simulation développé. Du fait d'une durée de simulation et d'une précision très dépendantes des valeurs des coefficients optiques, leurs variations lors des phases itératives d'optimisation posent des difficultés pour assurer convergence et robustesse au processus d'identification.

En amont, il est donc nécessaire d'explorer d'autres algorithmes de simulation du problème direct permettant de trouver un compromis durée – précision – complexité mieux adapté à une utilisation dans un schéma de résolution du problème inverse et en particulier des approches visant à limiter le nombre de photons lancés en s'intéressant d'abord aux parcours des photons collectés.

I.4. ANNEXES

I.4.1. Absorption, diffusion, fluorescence

Les interactions de la lumière avec la matière peuvent être décrite du point de vue ondulatoire (équations de Maxwell) ou du point de vue quantique (Equation de Transfert Radiatif). Cette seconde approche est favorisée dans le cas des tissus biologiques car elle permet de mieux prendre en compte la grande complexité de ces milieux (turbidité, hétérogénéité des formes et tailles des structures, absorption, diffusion et fluorescence) en utilisant des grandeurs moyennes caractéristiques du milieu.

Ainsi, on rappelle la définition de l'énergie d'un photon :

$$E_p = \frac{hc}{\lambda} \quad (6)$$

avec : h la constante de Planck, c la célérité de la lumière et λ la longueur d'onde.

Les propriétés photo-physiques des molécules font référence aux caractéristiques des transitions entre leurs différents niveaux ou états énergétiques probables (niveau d'énergie électronique, vibrationnel et rotationnel).

Lorsqu'un tissu biologique est exposé à un rayonnement lumineux dont la longueur d'onde est située dans la bande UV – Visible – Proche Infra-Rouge (350-850 nm), les interactions de ce rayonnement électromagnétique avec les molécules constitutives du tissu font essentiellement intervenir des transitions d'énergie électronique et peuvent être de différentes natures :

- phénomènes d'absorption

Une molécule dans son état fondamental (ou stable, ou de plus basse énergie) peut absorber l'énergie d'un photon incident et se trouver ainsi promue, très temporairement, à un état d'énergie supérieur instable. Pour certaines molécules, appelés chromophores, le retour de cet état excité de la molécule à son état fondamental s'effectue de façon non-radiative (agitation thermique). Pour d'autres molécules, appelés fluorophores, la transition radiative peut provoquer l'émission d'un photon d'énergie plus faible (donc de longueur d'onde supérieure). Les fluorophores endogènes sont des molécules fluorescentes naturellement présentes dans les tissus qui absorbent la lumière typiquement entre 300 et 450 nm et émettent une intensité de fluorescence intrinsèque (ou autofluorescence) sur une bande relativement large également.

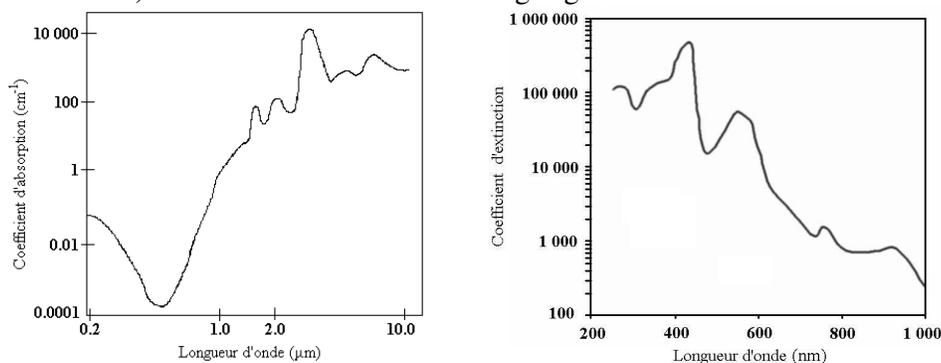


Figure 12. Spectres d'absorption de l'eau a) et d'extinction de l'hémoglobine b).

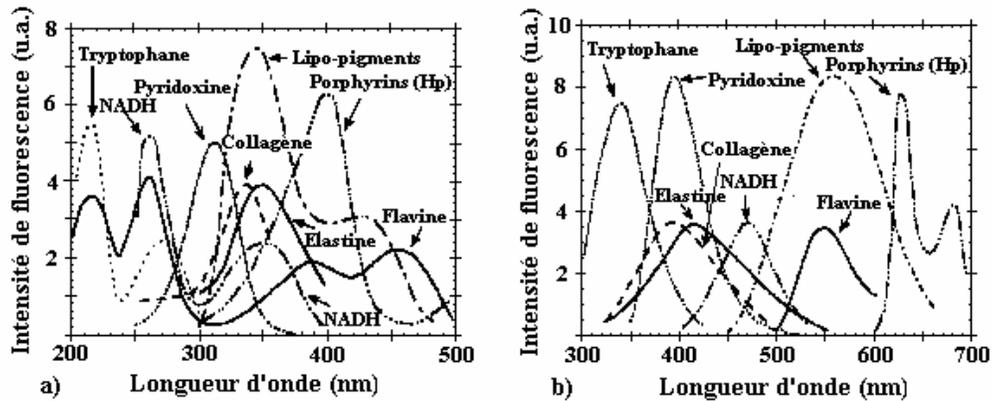


Figure 13. Spectres d'absorption a) et d'émission b) des fluorophores endogènes majeurs dans les tissus biologiques

- phénomènes de diffusion
Il existe différents phénomènes de diffusion mais nous évoquerons uniquement ici celui de la diffusion élastique qui est prépondérant aux longueurs d'ondes et aux puissances utilisées dans notre configuration. La diffusion élastique (ou collision élastique) est un processus instantané basé sur le phénomène physique d'interaction d'un photon avec une molécule au cours duquel l'énergie cinétique entre les deux particules est conservée (pas de perte d'énergie, pas de changement de longueur d'onde).

I.4.2. Spectroscopie de gap-FRAP et endoscopie de fluorescence multi-bandes

- Spectroscopie de gap-FRAP
La diminution du nombre et de la fonctionnalité des jonctions communicantes cellulaires (gap-junctions) intervient de façon précoce dans le processus de cancérisation. La technique appelée gap-FRAP (Fluorescence Recovery After Photobleaching) implémentée en microscopie et en spectroscopie, peut mettre en évidence des modifications spécifiques de l'activité de ces jonctions. Elle consiste à marquer les cellules du tissu à l'aide d'une sonde fluorescente puis à photoblanchir (détruire) avec un laser ceux situés dans une zone d'une ou plusieurs cellules et enfin à mesurer sur ce site le retour de l'intensité de fluorescence émise par les fluorophores des cellules adjacentes qui peuvent diffuser au travers des jonctions communicantes.
- Endoscopie de fluorescence multi-bandes
L'examen clinique visuel des organes creux pour la détection de lésions cancéreuses s'effectue classiquement grâce à l'endoscopie qui permet d'obtenir des images couleurs sous un éclairage en lumière blanche. Les systèmes d'imagerie endoscopique de fluorescence utilisent alternativement une source de lumière blanche et une source de lumière bleue permettant d'exciter les fluorophores endogènes. Ainsi, il est possible d'obtenir l'information contenue dans les spectres de fluorescence sous la forme d'une image de même taille et résolution que l'image en lumière blanche en endoscopie conventionnelle. L'image dite de fluorescence est formée à partir de la combinaison de plusieurs images acquises dans des bandes spectrales différentes. Un traitement d'images permet l'obtention d'une image en fausses couleurs reflétant la valeur des indicateurs spectraux caractéristiques du tissu. La visualisation simultanée et en temps réel de l'image conventionnelle en lumière blanche et de l'image de fluorescence correspondante doit permettre au clinicien d'explorer rapidement la totalité de la surface tissulaire à examiner et d'identifier les zones suspectes à biopsier.

I.5. REFERENCES BIBLIOGRAPHIQUES

W. Blondel, Ch. Daul, S. Villette, R. Miranda-Luna, Y. Hernandez-Mier, G. Bourg-Heckly and D. Wolf (2006). « Autofluorescence endoscopic spectro-imaging and 2D-cartography for *in situ* localisation and diagnosis of cancerous lesions. » *Traitement du Signal*, **23**: 213-233.

C. Choserot, E. Pery, J.-C. Goebel, D. Dumas, J. Didelon, J.-F. Stoltz and W.C.P.M. Blondel (2005). « Experimental comparison between autofluorescence spectra of constrained fresh and cryopreserved arteries » *Clinical Hemorheology and Microcirculation*, **33**: 235-42.

Q. Liu, C. Zhu, and N. Ramanujam (2003), « Experimental validation of Monte Carlo modeling of fluorescence in tissues in the UV-visible spectrum », *Journal of Biomedical Optics*, **8** (2): 223-236.

S.C. Pahl, « Light transport in tissue », PhD thesis, december 1998.

J. Prat, « Mesure des propriétés optiques de milieux diffusants stratifiés par l'analyse de la rétrodiffusion d'impulsions infrarouges subpicosecondes », PhD thesis, september 2002.

J. Selb, « Source virtuelle acousto-optique pour l'imagerie des milieux diffusants », PhD thesis, novembre 2002.

L. Voisin-Gobin, « Quantification de l'interaction lumière-tissus biologiques par la mesure non invasive du coefficient d'absorption et du coefficient réduit de diffusion », PhD thesis, mars 1999.

L Wang, S.L. Jacques, and L. Zheng (1995), « MCML – Monte Carlo modelling of light transport in multi-layered tissues », *Computer Methods and Programs in Biomedicine*, **47**: 131-146.