

La régression PLS 1, cas particulier de la régression linéaire séquentielle orthogonale (RLSO)

Jacques Goupy

ReConFor, 24 avenue Perrichont. 75016 Paris. France.

jacques@goupy.com

RÉSUMÉ

La régression linéaire classique fournit une seule solution souvent basée sur le critère des moindres carrés. Lorsqu'il y a beaucoup de variables, on risque d'obtenir un modèle surparamétré, c'est-à-dire modélisant les erreurs. Pour éviter cette surparamétrisation, la régression PLS a été introduite car étant une régression séquentielle, elle permet d'arrêter le processus de régression avant de modéliser l'erreur. Mais la PLS ne propose que quelques solutions et ne permet pas vraiment d'optimisation. Nous proposons dans cet article une régression séquentielle orthogonale beaucoup plus souple que la PLS, la régression linéaire séquentielle orthogonale ou RLSO. Cette régression permet d'éviter la surparamétrisation et elle facilite l'optimisation de la modélisation car les directions de projection peuvent être choisies au mieux des intérêts de l'étude. Les solutions possibles proposées par la RLSO sont en nombre infini. Ce qui laisse la possibilité d'utiliser toutes les techniques d'optimisation pour trouver la ou les meilleures solutions.

Comme la régression linéaire séquentielle orthogonale (RLSO) est nouvelle, nous avons développé toutes les étapes de calcul.

Mots clés: Régression séquentielle, Régression PLS1, Régression orthogonale.

ABSTRACT

The classical linear regression provides just one solution often based on the least squares criterion. When there are many variables, one can obtain a model with too many surparameters and consequently a model which modelizes the error. To avoid such an over parameterization, the PLS regression has been introduced. PLS is a sequential regression and it enables to stop the regression process before the error modelling. But the PLS regression has only a few numbers of solutions and the optimization is not really possible. It is why a new type of sequential regression is proposed in this article: the sequential orthogonal linear regression or SOL regression. It is much more versatile than the PLS ; it enables to avoid surparametrization and facilitate the modelling optimization because the directions of projection can be chosen at the best for the study. The number of the solutions proposed by the SOL regression is infinite and all the techniques of optimization can be used.

As the SOL regression is new, all the calculations have been largely explained.

Key words: sequential regression, PLS1 regression, orthogonal regression.

1. INTRODUCTION

La régression linéaire a pour objectif de trouver la relation mathématique linéaire qui représente le mieux le lien entre une grandeur d'intérêt (ou réponse), y , et des grandeurs explicatives (ou facteurs), x . La régression possède une terminologie très riche et très diverse pour désigner les grandeurs d'intérêt et les grandeurs explicatives (Tableau 1).

Grandeur d'intérêt y	Variable x
Variable expliquée Variable endogène Variable dépendante Réponse Observation Variable d'intérêt	Variable explicative Variable exogène Variable indépendante Facteur Régresseur Prédicteur Contrôle

Tableau 1 : Principales terminologies des grandeurs d'intérêt et des variables

Dans cet article nous avons choisi la terminologie des plans d'expériences c'est-à-dire "réponse" pour grandeur d'intérêt et "facteurs" pour variables explicatives. Ces facteurs ont des niveaux, x_i , définis ou mesurés par l'observateur. On introduit un écart, e , pour tenir compte du fait que le modèle retenu n'est pas exactement le modèle réel et que la grandeur d'intérêt est une grandeur aléatoire. On a donc

$$y = f(x_1, x_2, \dots, x_n) + e$$

Dans la régression linéaire, on suppose que la fonction f est une fonction linéaire. Pour établir cette fonction, c'est-à-dire trouver la valeur des coefficients du modèle, il faut résoudre un système d'équations linéaires dont les termes connus sont les y et les x_i , et dont les termes inconnus sont les coefficients et les écarts. Nous supposons que toutes les grandeurs sont centrées et réduites, ce qui ne retire aucune généralité à la théorie ou aux calculs. On écrit le système à résoudre sous forme matricielle

$$\mathbf{y} = \mathbf{X} \mathbf{a} + \mathbf{e} \quad \{1\}$$

où

- \mathbf{y} est le vecteur $(n,1)$ des réponses mesurées ou observées. C'est le vecteur à régresser. Cette matrice-vecteur est centrée et réduite.
- \mathbf{X} est la matrice (n,p) des niveaux des facteurs et du modèle mathématique. Cette matrice est centrée et réduite. Elle est supposée de plein rang.
- \mathbf{a} est le vecteur $(p,1)$ des coefficients. C'est le vecteur à déterminer.
- \mathbf{e} est le vecteur $(n,1)$ des écarts. Cette matrice est inconnue.

S'il y a n réponses y et p coefficients dans le modèle, il y a n équations et $n + p$ inconnues. Pour trouver les p équations manquantes, on utilise la méthode des moindres carrés qui consiste à minimiser la somme des carrés des écarts. Le système de la régression linéaire classique est donc

$$\begin{cases} \mathbf{y} = \mathbf{X} \mathbf{a} + \mathbf{e} \\ \frac{\partial {}^t \mathbf{e} \mathbf{e}}{\partial \mathbf{a}} = 0 \end{cases} \quad \{2\}$$

En faisant appel à quelques hypothèses simplificatrices, on obtient la solution de ce système, c'est-à-dire que l'on obtient les coefficients :

$$\hat{\mathbf{a}} = ({}^t \mathbf{X} \mathbf{X})^{-1} {}^t \mathbf{X} \mathbf{y} \quad \{3\}$$

Ayant ces coefficients, on a la possibilité d'obtenir les réponses calculées connaissant les niveaux des facteurs :

$$\hat{\mathbf{y}} = \mathbf{X} \hat{\mathbf{a}} \quad \{4\}$$

où

- $\hat{\mathbf{y}}$ est le vecteur $(n,1)$ des réponses calculées avec la méthode des moindres carrés. C'est la solution classique des moindres carrés.
- $\hat{\mathbf{a}}$ est le vecteur $(p,1)$ des coefficients calculés avec la méthode des moindres carrés.

A partir de ces résultats, on peut écrire plusieurs relations qui sont fort utiles pour obtenir les valeurs numériques des différentes grandeurs. On introduit la matrice de projection \mathbf{H} , ou matrice "hat", telle que :

$$\mathbf{H} = \mathbf{X} [{}^t \mathbf{X} \mathbf{X}]^{-1} {}^t \mathbf{X} \quad \{5\}$$

La relation {4} peut alors s'écrire

$$\hat{\mathbf{y}} = \mathbf{H} \mathbf{y} \quad \{6\}$$

La matrice \mathbf{H} associée à la matrice unité \mathbf{I} permet également de calculer le vecteur des résidus.

$$\hat{\mathbf{e}} = (\mathbf{I} - \mathbf{H}) \mathbf{y} \quad \{7\}$$

Les coefficients du modèle peuvent également être exprimés par la relation

$$\hat{\mathbf{a}} = ({}^t \mathbf{X} \mathbf{X})^{-1} {}^t \mathbf{X} \hat{\mathbf{y}} \quad \{8\}$$

Ces formules matricielles peuvent être illustrées par des représentations géométriques donnant ainsi un éclairage intéressant de la régression ainsi qu'un support visuel à la réflexion et à l'analyse.

2. LES CONCEPTS

Supposons que le modèle ne comporte qu'un coefficient et que deux mesures de la réponse aient été réalisées, l'une au niveau $x_{1,1}$ et l'autre au niveau $x_{1,2}$. Les valeurs des réponses sont respectivement y_1 et y_2 . Le système d'équations est :

$$\begin{cases} y_1 = a_1 x_{1,1} + e_1 \\ y_2 = a_1 x_{1,2} + e_2 \end{cases}$$

On peut écrire ce système d'équations sous forme matricielle pour faire apparaître les vecteurs :

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_{1,1} \\ x_{1,2} \end{bmatrix} [a_1] + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

avec

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} x_{1,1} \\ x_{1,2} \end{bmatrix} \quad \mathbf{a} = [a_1] \quad \mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

Pour illustrer cette relation, introduisons un référentiel dont le premier axe porte la réponse y_1 et dont le second, orthogonal au premier, porte la réponse y_2 . Ce référentiel sera appelé « référentiel des réponses ». On peut tracer le vecteur \mathbf{y} des réponses mesurées (Figure 1) qui a pour composantes y_1 et y_2 .

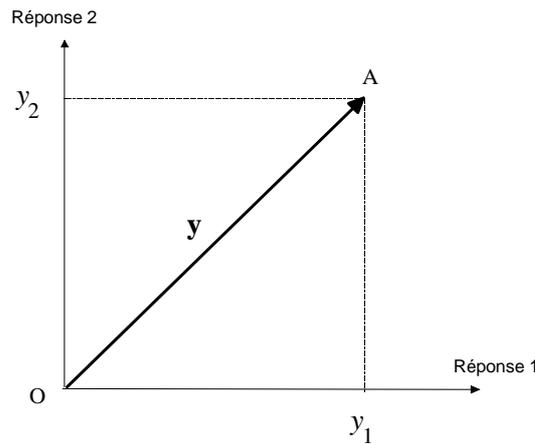


Figure 1 : Le vecteur \overrightarrow{OA} représente les réponses mesurées dans le référentiel des réponses.

Le vecteur \mathbf{y} est la somme de deux vecteurs, le vecteur des réponses possibles $\begin{bmatrix} x_{1,1} \\ x_{1,2} \end{bmatrix} [a_1]$ et le vecteur des écarts $\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$.

Les niveaux $x_{1,1}$ et $x_{1,2}$ sont constants et parfaitement définis selon les hypothèses de la régression. Le rapport de ces niveaux est également constant :

$$\frac{x_{1,1}}{x_{1,2}} = \text{constant}$$

La valeur du coefficient a_1 est inconnue et c'est elle que l'on cherche à déterminer. Si l'on donne la valeur unité à ce coefficient, on détermine le point M_1 de coordonnées $x_{1,1}$ et $x_{1,2}$ (Figure 2). Si l'on donne la valeur 2 à ce coefficient, on détermine un point M_2 de coordonnées $2x_{1,1}$ et $2x_{1,2}$. Si l'on donne la valeur zéro à ce coefficient, on obtient le point O, origine du référentiel.

Si l'on fait varier la valeur de ce coefficient de $-\infty$ à $+\infty$, le point N parcourt une droite D. Cette droite peut être graduée en fonction des valeurs du coefficient a_1 . Chaque point de la droite D représente une solution possible du système d'équations, c'est pourquoi cette droite s'appelle « l'espace des réponses calculées possibles » ou plus simplement « l'espace des réponses possibles » ou ERP.

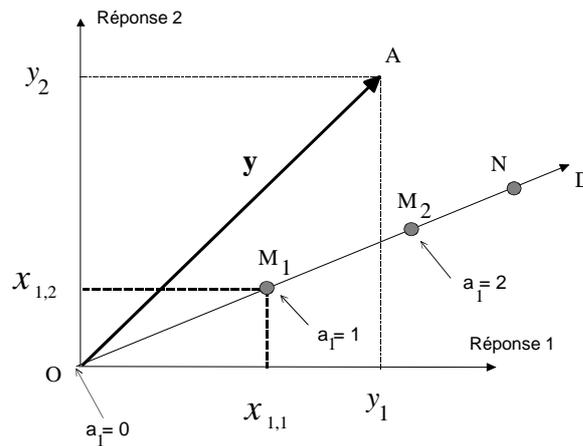


Figure 2 : La droite D représente l'ensemble de toutes les réponses calculées possibles.

On appellera "vecteur niveau", le vecteur $\overrightarrow{OM_1}$ dont les composantes sont $x_{1,1}$ et $x_{1,2}$. C'est ce vecteur qui définit la direction de la droite D . C'est donc aussi le vecteur niveau qui définit l'ERP.

$$\mathbf{x}_1 = \overrightarrow{OM_1} = \begin{bmatrix} x_{1,1} \\ x_{1,2} \end{bmatrix}$$

Le vecteur \overrightarrow{OA} est égal au vecteur \overrightarrow{ON} augmenté du vecteur \overrightarrow{NA} (Figure 3), on a

$$\overrightarrow{OA} = \overrightarrow{ON} + \overrightarrow{NA}$$

Le vecteur \overrightarrow{OA} est une représentation de la matrice \mathbf{y} des réponses mesurées. Le vecteur \overrightarrow{ON} est une représentation de la matrice des réponses calculées possibles, \mathbf{y}' , et le vecteur \overrightarrow{NA} est une représentation de la matrice \mathbf{e} des écarts.

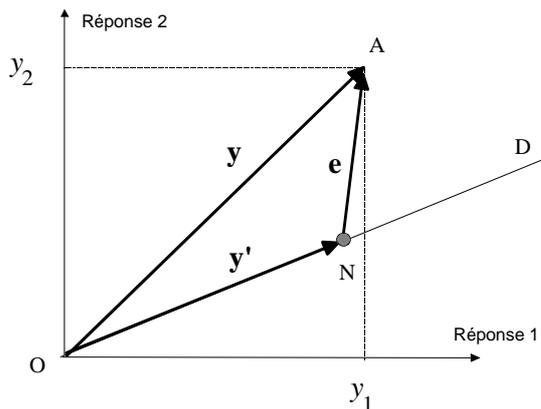


Figure 3 : Le vecteur des réponses mesurées est la somme vectorielle du vecteur des réponses possibles et du vecteur des écarts.

Dans le cadre du critère des moindres carrés, on suppose que la solution la plus vraisemblable est donnée par le vecteur des écarts le plus petit possible. On obtient cette solution lorsque le vecteur \mathbf{e} est orthogonal à la droite D . La solution est

alors le vecteur que nous avons noté \hat{y} . Le vecteur des écarts devient le vecteur des résidus noté \hat{e} . Le point N prend la position particulière H (Figure 4).

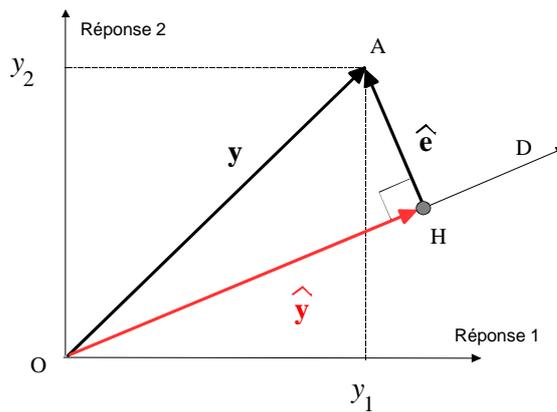


Figure 4 : Le vecteur des résidus, \hat{e} , est orthogonal à la droite D des réponses possibles ainsi qu'au vecteur des réponses calculées \hat{y} .

Il est important de noter, que dans le cadre de la théorie des moindres carrés, le vecteur des résidus \hat{e} est orthogonal au vecteur des réponses calculées, \hat{y} . On peut dire aussi que le vecteur des réponses calculées, \hat{y} , est la projection orthogonale sur la droite D du vecteur des réponses mesurées, y . On peut dire aussi que régresser un ensemble de données, c'est projeter orthogonalement le vecteur représentatif de la matrice des réponses mesurées sur l'espace des réponses possibles.

Les trois côtés du triangle rectangle OHA jouent un rôle essentiel dans la théorie de la régression et nous les retrouverons constamment :

- Le vecteur y des réponses mesurées ou vecteur à régresser est l'hypoténuse du triangle rectangle OHA.
- Le vecteur \hat{y} ou vecteur régressé est le côté de l'angle droit qui est dans l'ERP.
- Le vecteur \hat{e} ou vecteur des résidus est le côté de l'angle droit qui n'est pas dans l'ERP.

La relation {4}, $\hat{y} = \mathbf{X} \hat{a}$, peut être interprétée de la manière suivante : Le vecteur \overrightarrow{OH} représente aussi bien le vecteur \hat{a} que le vecteur \hat{y} . Le vecteur \hat{a} est défini dans l'ERP et possède 1 composante. Le vecteur \hat{y} est défini dans le référentiel des réponses et il possède 2 composantes. La matrice \mathbf{X} transforme donc un vecteur de l'ERP en son équivalent dans le référentiel des réponses.

3. LA REGRESSION LINEAIRE CLASSIQUE

Les concepts introduits à l'aide de deux réponses et d'un coefficient peuvent facilement être étendus à n réponses et à p coefficients.

Le référentiel des réponses possède n dimensions et c'est un référentiel orthonormé. Le vecteur des réponses mesurées possède également n dimensions.

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}$$

On peut définir p droites D engendrant l'ERP qui possède donc p dimensions. Les directions de ces p droites sont définies par les vecteurs niveaux, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$. Ces vecteurs n'étant généralement ni orthogonaux ni de mêmes normes, l'ERP est défini par un référentiel oblique non normé. Les droites D sont graduées en fonction des valeurs des coefficients correspondants et chaque point de l'ERP représente un jeu possible de coefficients pour le modèle de régression, donc une solution possible.

Les vecteurs $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ sont les colonnes de la matrice \mathbf{X} . L'espace des réponses possibles est donc défini par les colonnes de la matrice \mathbf{X} .

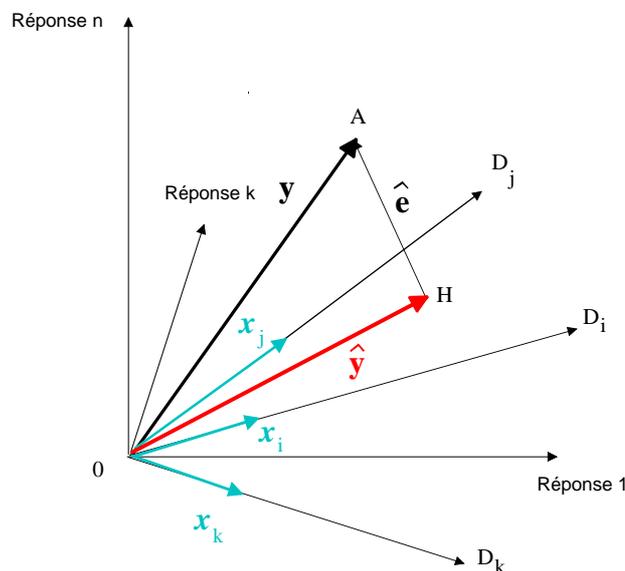


Figure 5 : Le vecteur des réponses mesurées, \mathbf{y} , est décomposé en deux vecteurs : le vecteur des réponses calculées, $\hat{\mathbf{y}}$ et le vecteur des résidus, $\hat{\mathbf{e}}$.

La régression classique consiste à trouver le vecteur des écarts ayant la norme la plus petite possible. Comme dans le paragraphe précédent, c'est la projection orthogonale de \mathbf{y} sur l'ERP qui donne la solution classique des moindres carrés. On obtient ainsi le vecteur des réponses calculées, $\hat{\mathbf{y}}$, et le vecteur des résidus, $\hat{\mathbf{e}}$. Les vecteurs $\hat{\mathbf{y}}$ et $\hat{\mathbf{e}}$ sont orthogonaux (Figure 5).

4. REGRESSION LINEAIRE SUR UNE DIRECTION

Au lieu de régresser le vecteur \mathbf{y} sur l'hyperplan de l'EPR, on peut choisir une direction particulière dans cet espace. Choisissons une direction, Δ , située dans l'espace des réponses possibles. Cette direction est définie par son vecteur unitaire \mathbf{t} . Le vecteur \mathbf{t} est repéré par ses paramètres directeurs t_i .

$$\mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_p \end{bmatrix}$$

L'ERP ayant p dimensions, le vecteur unitaire \mathbf{t} possède p paramètres directeurs t_i .

Régresser le vecteur \mathbf{y} sur la droite Δ , c'est le projeter orthogonalement sur cette droite. La projection de \mathbf{y} sur Δ est un vecteur de composantes \mathbf{b} que l'on cherche à déterminer. Ces composantes sont égales, à un coefficient de proportionnalité près, aux paramètres directeurs de la direction choisie sur laquelle on va projeter :

$$\mathbf{b} = k \mathbf{t}$$

Le facteur de proportionnalité n'est pas connu. On le calcule à l'aide de la méthode des moindres carrés. Il faut donc résoudre le système d'équations suivant :

$$\begin{cases} \mathbf{y} = \mathbf{X} \mathbf{b} + \mathbf{e} \\ \frac{\partial {}^t \mathbf{e} \mathbf{e}}{\partial k} = 0 \\ \mathbf{b} = k \mathbf{t} \end{cases} \quad \{9\}$$

- \mathbf{y} matrice $(n,1)$ des réponses mesurées. C'est le vecteur à régresser.
- \mathbf{X} matrice (n,p) des niveaux des facteurs et du modèle mathématique.
- \mathbf{b} matrice $(p,1)$ des coefficients. Ces coefficients sont proportionnels aux paramètres directeurs de \mathbf{t} .
- \mathbf{t} matrice $(p,1)$ des paramètres directeurs de la droite sur laquelle on projette. Ces paramètres sont connus puisque choisis par l'expérimentateur.
- \mathbf{e} matrice $(n,1)$ des écarts.
- k coefficient de proportionnalité. C'est l'inconnue à déterminer.

La solution des moindres carrés est la suivante (voir détail du calcul en annexe 1) :

$$\tilde{k} = [{}^t \mathbf{t} {}^t \mathbf{X} \mathbf{X} \mathbf{t}]^{-1} {}^t \mathbf{t} {}^t \mathbf{X} \mathbf{y} \quad \{10\}$$

D'où (Figure 6)

- les composantes du vecteur régressé : $\tilde{\mathbf{b}} = \tilde{k} \mathbf{t}$
- le vecteur projeté sur D, solution de la régression $\tilde{\mathbf{y}} = \mathbf{X} \tilde{\mathbf{b}} = \tilde{k} \mathbf{X} \mathbf{t}$

La relation, $\tilde{\mathbf{y}} = \mathbf{X} \tilde{\mathbf{b}}$, peut être interprétée de la manière suivante : Les vecteurs $\tilde{\mathbf{b}}$ et \mathbf{t} sont définis dans l'ERP, ils ont donc p composantes. Le vecteur $\tilde{\mathbf{y}}$ est défini dans le référentiel des réponses, il possède donc n composantes. La matrice \mathbf{X} transforme le vecteur $\tilde{\mathbf{b}}$ (ou $\tilde{k} \mathbf{t}$) de l'ERP en un vecteur équivalent, $\tilde{\mathbf{y}}$, du référentiel des réponses. La représentation graphique ne différencie pas ces deux vecteurs qui sont parfaitement superposés. La matrice \mathbf{X} transforme un vecteur de p composantes en un vecteur de n composantes.

Au vecteur \mathbf{t} (p composantes) défini dans l'ERP correspond le transformé \mathbf{Xt} qui est repéré dans le référentiel des réponses (n composantes). Le référentiel des réponses étant orthonormé, nous effectuerons les calculs dans ce repère.

- le vecteur des résidus de la régression $\mathbf{e} = \mathbf{y} - \tilde{\mathbf{y}}$

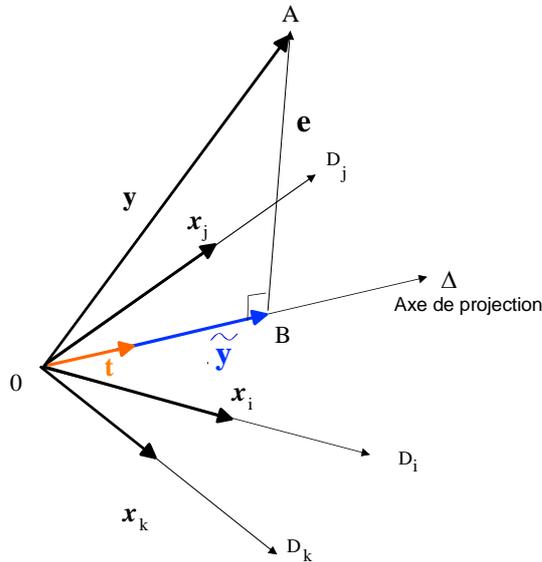


Figure 6 : Régression du vecteur \mathbf{y} sur la droite Δ de paramètres directeurs \mathbf{t} .

Le triangle rectangle OBA est rectangle. On retrouve :

Le vecteur à régresser : \mathbf{y}

Le vecteur régressé $\tilde{\mathbf{y}}$

Le vecteur des résidus \mathbf{e}

On peut également introduire la matrice de projection \mathbf{H}_Δ sur Δ :

$$\mathbf{H}_\Delta = \mathbf{Xt} [\mathbf{t}' \mathbf{X} \mathbf{X} \mathbf{t}]^{-1} \mathbf{t}' \mathbf{X} \quad \{11\}$$

On a le vecteur régressé

$$\tilde{\mathbf{y}} = \mathbf{H}_\Delta \mathbf{y} \quad \{12\}$$

et le vecteur des résidus

$$\mathbf{e} = (\mathbf{I} - \mathbf{H}_\Delta) \mathbf{y} \quad \{13\}$$

Les coefficients du modèle correspondant au vecteur $\tilde{\mathbf{y}}$ sont donnés par

$$\tilde{\mathbf{b}} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \tilde{\mathbf{y}} \quad \{14\}$$

La matrice $(\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}'$ transforme le vecteur $\tilde{\mathbf{y}}$ de n composantes en son équivalent, le vecteur $\tilde{\mathbf{b}}$ à p composantes.

5. REGRESSION LINEAIRE SEQUENTIELLE ORTHOGONALE

Au lieu d'obtenir immédiatement la solution des moindres carrés comme dans la régression classique, on peut procéder par étapes. On projette d'abord sur une direction quelconque de l'ERP puis, on cherche à retrouver progressivement le vecteur \hat{y} . Nous prendrons d'abord le cas d'un modèle à deux coefficients, l'EPR ne possède alors que deux dimensions et nous verrons que le vecteur régressé de la régression classique, \hat{y} , est décomposé en deux vecteurs orthogonaux.

Puis, nous étendrons notre étude à un modèle à trois coefficients. Le vecteur \hat{y} est alors décomposé en trois vecteurs orthogonaux. Enfin, nous verrons le cas du modèle à n coefficients et d'une décomposition de \hat{y} en n vecteurs orthogonaux.

5.1 DEUX COEFFICIENTS

Dans le cas d'un modèle à deux coefficients, on a deux vecteurs niveaux, x_1 et x_2 , portés par 2 droites D_1 et D_2 . Ces 2 droites définissent l'ERP1. La régression classique de y conduit au vecteur \hat{y} situé dans l'ERP1 (Figure 7).

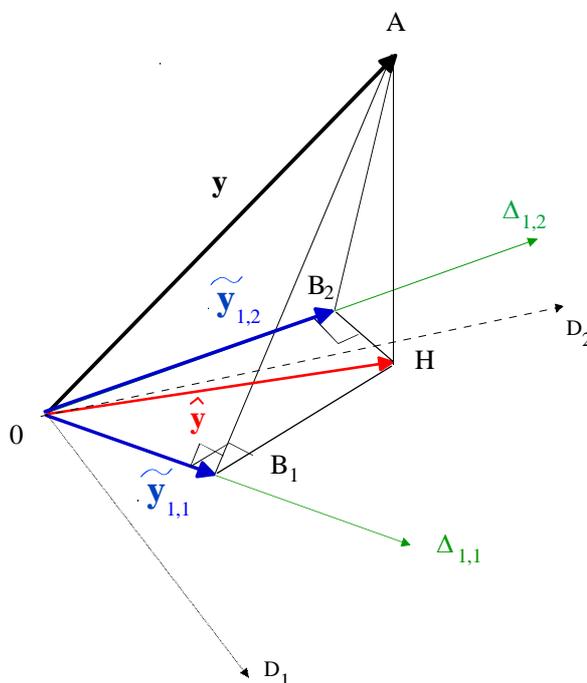


Figure 7 : Le vecteur y est régressé sur la droite $\Delta_{1,1}$, ou sur la droite $\Delta_{1,2}$, ou sur une droite $\Delta_{1,i}$ du plan $D_1 D_2$.

On projette y sur une première direction $\Delta_{1,1}$. On obtient un premier vecteur des réponses calculées, $\overrightarrow{OB_1} = \tilde{y}_{1,1}$. Les triangles OB_1H et OB_1A sont rectangles en B_1 .

On projette y sur une deuxième direction $\Delta_{1,2}$. On obtient un deuxième vecteur des réponses calculées, $\overrightarrow{OB_2} = \tilde{y}_{1,2}$. Les triangles OB_2H et OB_2A sont rectangles en B_2 . On pourrait projeter y sur n'importe quelle autre direction $\Delta_{1,i}$ de l'ERP1.

Les angles droits OB_1H et OB_2H sont sous tendus par le même segment OH qui n'est autre que le vecteur \hat{y} . Les points B_1 et B_2 sont donc sur un cercle de diamètre OH (Figure 8) et l'on a

$$\hat{y} = \tilde{y}_{1,1} + \overrightarrow{B_1H} \quad \text{ou} \quad \hat{y} = \tilde{y}_{1,2} + \overrightarrow{B_2H}$$

D'une manière générale, quel que soit le point B_i , on a

$$\hat{y} = \tilde{y}_{1,i} + \overrightarrow{B_iH}$$

C'est-à-dire que le vecteur \hat{y} peut être décomposé en deux vecteurs orthogonaux, $\tilde{y}_{1,i}$ et $\overrightarrow{B_iH}$.

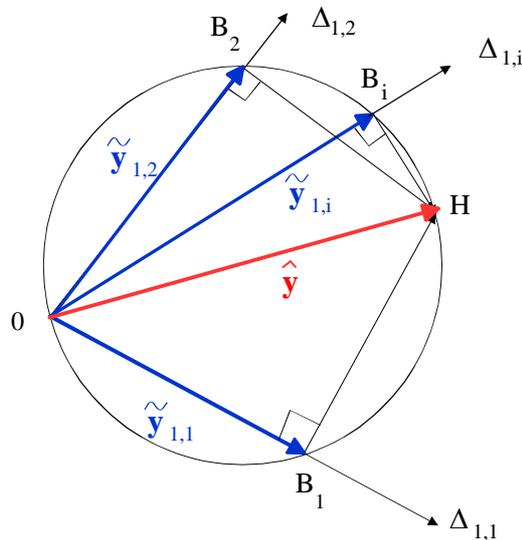


Figure 8 : le vecteur \hat{y} est décomposé en deux vecteurs orthogonaux.

Choisissons une direction particulière Δ_1 pour projeter le vecteur y . On obtient le vecteur régressé, \tilde{y}_1 , et le vecteur des résidus e_1 . Régressons maintenant le vecteur e_1 sur une direction Δ_2 orthogonale à Δ_1 et définit par BH. On obtient le vecteur régressé, \tilde{e}_2 , et le vecteur des résidus e_2 (Figure 9).

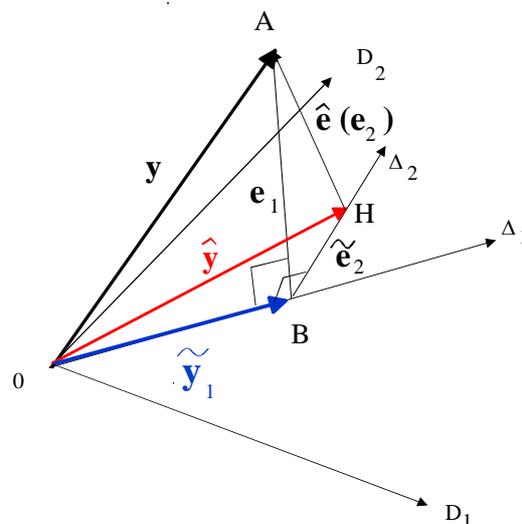


Figure 9 : Le vecteur e_1 est régressé sur la direction Δ_2 parallèle à la droite BH.

Examinons le triangle AHB rectangle en H. On constate que la régression de \mathbf{e}_1 sur la direction du vecteur BH conduit au vecteur régressé BH lui-même et que le vecteur des résidus, \mathbf{e}_2 , n'est autre que le vecteur des résidus, $\hat{\mathbf{e}}$, de la régression classique.

La Figure 10 extrait deux triangles importants de la Figure 9 pour faciliter la compréhension :

- le triangle OBH qui rassemble le vecteur projeté de la régression classique, $\hat{\mathbf{y}}$; le vecteur projeté sur la direction Δ_1 , $\tilde{\mathbf{y}}_1$; le vecteur $\tilde{\mathbf{e}}_2$ (régressé de \mathbf{e}_1) orthogonal à Δ_1 .
- le triangle AHB qui rassemble le vecteur des résidus de la régression de \mathbf{y} sur la direction Δ_1 , \mathbf{e}_1 ; le vecteur des résidus de la régression classique, $\hat{\mathbf{e}}$; le vecteur régressé de la régression de \mathbf{e}_1 sur la direction BH, $\tilde{\mathbf{e}}_2$.

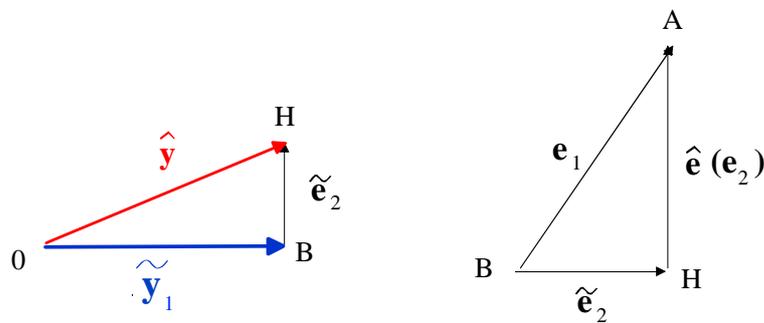


Figure 10 : Triangles OBH et AHB extraits de la figure 9.

Pour réaliser la régression de \mathbf{e}_1 sur la direction de BH, il faut connaître les paramètres directeurs de cette direction. Pour cela, on régresse les vecteurs niveaux \mathbf{x}_1 et \mathbf{x}_2 sur Δ_1 . On obtient deux vecteurs des résidus $\mathbf{r}_{1,1}$ et $\mathbf{r}_{1,2}$ qui sont orthogonaux à Δ_1 et qui sont donc parallèles à BH puisque l'on opère dans un plan, le plan $D_1 D_2$ (Figure 11). On a ainsi les paramètres directeurs (identiques à ceux de BH) de la direction Δ_2 , orthogonale à Δ_1 .

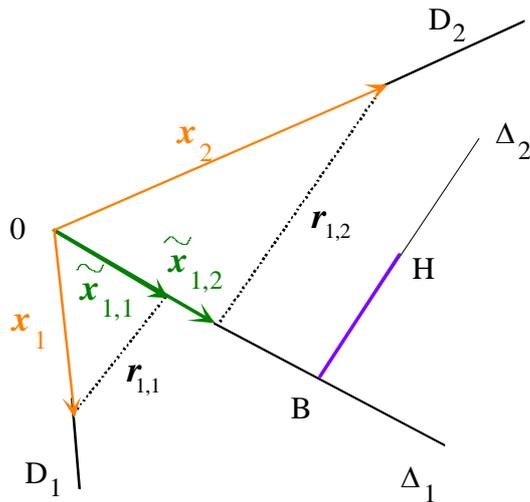


Figure 11 : Les vecteurs niveaux x_1 et x_2 sont décomposés en deux vecteurs projetés $\tilde{x}_{1,1}$ et $\tilde{x}_{1,2}$ et deux vecteurs des résidus $r_{1,1}$ et $r_{1,2}$

Finalement, pour retrouver \hat{y} en effectuant une régression séquentielle orthogonale, il faut réaliser deux projections successives sur deux directions orthogonales : Δ_1 et Δ_2 .

Analysons ces deux projections.

Projections sur Δ_1

Nous allons considérer deux projections (ou régressions) sur la droite Δ_1 de paramètres directeurs t_1 . La première projection correspond à la régression du vecteur y qui fournit une première solution \tilde{y}_1 . La seconde projection correspond à la construction d'un nouvel ERP orthogonal à la droite Δ_1 .

Régression de y sur Δ_1

Nous avons déjà indiqué les relations donnant le vecteur régressé \tilde{y}_1 , sur la droite Δ_1 et le vecteur des résidus e_1 correspondant :

$$\tilde{y}_1 = H_1 y \quad \{15\}$$

$$e_1 = (I - H_1) y \quad \{16\}$$

avec la matrice de projection sur Δ_1

$$H_1 = X t_1 [t_1' X X t_1]^{-1} t_1' X \quad \{17\}$$

Le choix de cette première direction de projection est entièrement libre et peut être déterminé au mieux des intérêts de l'étude.

La matrice $(I - H_1)$ est la matrice de projection sur un espace orthogonal à Δ_1 .

Régression de x_1 et x_2 sur Δ_1

Cette régression a pour objectif d'obtenir un ERP orthogonal à Δ_1 et par conséquent, dans le cas de deux coefficients, parallèle à BH . Lorsqu'on régresse les vecteurs x_1 et x_2 sur Δ_1 , on obtient les vecteurs régressés $\tilde{x}_{1,1}$ et $\tilde{x}_{1,2}$ qui sont portés

par la droite Δ_1 et les vecteurs des résidus $r_{1,1}$ et $r_{1,2}$ qui sont orthogonaux à cette même droite (Figure 12). On a

$$\begin{aligned}\tilde{x}_{1,1} &= \mathbf{H}_1 x_1 & r_{1,1} &= (\mathbf{I} - \mathbf{H}_1) x_1 \\ \tilde{x}_{1,2} &= \mathbf{H}_1 x_2 & r_{1,2} &= (\mathbf{I} - \mathbf{H}_1) x_2\end{aligned}$$

Le nouvel espace des réponses possibles est orthogonal à Δ_1 et son origine est en B. Il a une dimension de moins que ERP1, et par conséquent, il n'a qu'une dimension. C'est donc une droite, la droite Δ_2 . L'ERP2 est engendré par les vecteurs des résidus $r_{1,1}$ et $r_{1,2}$. Ces vecteurs peuvent être regroupés dans une matrice \mathbf{X}_{r_1} telle que

$$\mathbf{X}_{r_1} = [r_{1,1} \quad r_{1,2}]$$

Cette matrice permet d'écrire la relation suivante qui montre que l'ERP1 est projeté sur un espace orthogonal à Δ_1 par la matrice de projection $(\mathbf{I} - \mathbf{H}_1)$:

$$\mathbf{X}_{r_1} = (\mathbf{I} - \mathbf{H}_1) \mathbf{X} \quad \{18\}$$

Les vecteurs des résidus $r_{1,1}$ et $r_{1,2}$ sont dans le plan défini par les droites D_1 et D_2 , étant orthogonaux à la droite Δ_1 , ils sont parallèles entre eux. Ils ne sont donc pas linéairement indépendants et la matrice \mathbf{X}_{r_1} n'est pas de plein rang.

Au vecteur t_1 (2 composantes) défini dans l'ERP1 correspond le transformé $\mathbf{X} t_1$ qui est repéré dans le référentiel des réponses et qui possède n composantes.

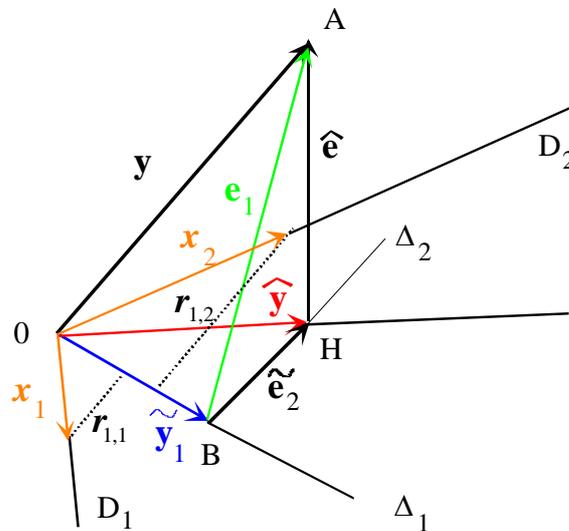


Figure 12 : Régession des vecteurs x_1 et x_2 sur la direction de la droite Δ_2 .

Projections sur Δ_2

On considère une projection (ou régression) sur la droite Δ_2 appartenant à ERP2 et dont les paramètres directeurs t_2 sont définis dans l'ERP1. Cette projection correspond à la décomposition du vecteur des résidus e_1 et elle conduit à la seconde solution \tilde{y}_2 de la régression qui, dans ce cas, est \hat{y} .

Le choix des paramètres directeurs de t_2 est entièrement libre.

Régression de \mathbf{e}_1 sur la direction Δ_2

Le vecteur BH peut être regardé comme le régressé de \mathbf{e}_1 sur la direction de mêmes paramètres directeurs que les vecteurs $\mathbf{r}_{1,1}$ et $\mathbf{r}_{1,2}$. Cette régression s'effectue dans un plan orthogonal à Δ_1 en B. Le résidu de cette régression est le résidu $\hat{\mathbf{e}}$ de la régression classique de \mathbf{y} .

On obtient le vecteur régressé :

$$\overrightarrow{\text{BH}} = \tilde{\mathbf{e}}_2 = \mathbf{H}_2 \mathbf{e}_1$$

avec

$$\mathbf{H}_2 = \mathbf{X}_{r_1} \mathbf{t}_2 \left[\mathbf{t}_2' \mathbf{X}_{r_1} \mathbf{X}_{r_1} \mathbf{t}_2 \right]^{-1} \mathbf{t}_2' \mathbf{X}_{r_1} \quad \{19\}$$

Au vecteur \mathbf{t}_2 de l'ERP1 correspond le vecteur $\mathbf{X}_{r_1} \mathbf{t}_2$ (n composantes) défini dans le référentiel des réponses.

Réduction dimensionnelle des ERP

La projection $(\mathbf{I} - \mathbf{H}_1)$ projette les vecteurs de l'ERP1 sur une droite orthogonale à Δ_1 (Figure 13). Elle transforme le vecteur \mathbf{t}_2 défini dans ERP1, en un vecteur \mathbf{t}'_2 orthogonal à Δ_1 situé dans l'ERP2 qui ne possède qu'une seule dimension. La droite Δ_2 est parallèle à \mathbf{t}'_2 .

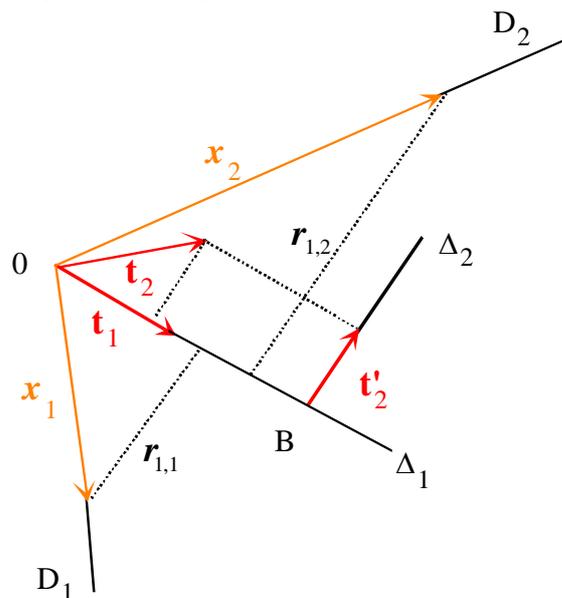


Figure 13 : La projection $(\mathbf{I} - \mathbf{H}_1)$ transforme l'ERP1 ayant deux dimensions en ERP2 n'ayant plus qu'une dimension.

Les solutions de la régression

Il y a deux vecteurs pouvant prétendre être solution de la régression : le vecteur $\tilde{\mathbf{y}}_1$ projeté sur Δ_1 et le vecteur $\hat{\mathbf{y}}$ de la solution classique (Figure 14). On a

$$\tilde{\mathbf{y}}_1 = \mathbf{H}_1 \mathbf{y}$$

$$\hat{\mathbf{y}} = \tilde{\mathbf{y}}_1 + \tilde{\mathbf{e}}_2 = \tilde{\mathbf{y}}_1 + \mathbf{H}_2 \mathbf{e}_1$$

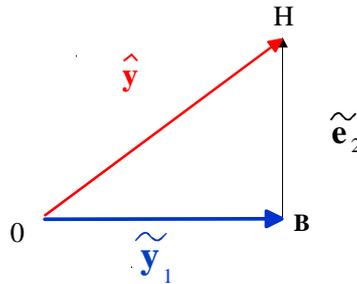


Figure 14 : Illustration de la décomposition du vecteur $\hat{\mathbf{y}}$.

Coefficients des vecteurs régressés

Les coefficients $\tilde{\mathbf{b}}_1$ du modèle mathématique donnant le vecteur régressé $\tilde{\mathbf{y}}_1$, sont fournis par la relation

$$\tilde{\mathbf{b}}_1 = \tilde{k} \mathbf{t}_1 = \mathbf{t}_1 \left[\mathbf{t}_1' \mathbf{X} \mathbf{X} \mathbf{t}_1 \right]^{-1} \mathbf{t}_1' \mathbf{X} \mathbf{y}$$

Les coefficients s'obtiennent également avec la relation :

$$\tilde{\mathbf{b}}_1 = (\mathbf{t}_1' \mathbf{X} \mathbf{X} \mathbf{t}_1)^{-1} \mathbf{t}_1' \mathbf{X} \tilde{\mathbf{y}}_1$$

Les coefficients $\hat{\mathbf{a}}$ du modèle mathématique donnant le vecteur régressé $\hat{\mathbf{y}}$, sont fournis par la relation classique:

$$\hat{\mathbf{a}} = \left[\mathbf{X}' \mathbf{X} \right]^{-1} \mathbf{X}' \mathbf{y}$$

Résumé des principales opérations

Le Tableau 2 et le Tableau 3 indiquent les principales opérations effectuées dans ERP1 et ERP2.

	Projection sur Δ_1	Projection \perp à Δ_1
	\mathbf{H}_1	$(\mathbf{I} - \mathbf{H}_1)$
\mathbf{y}	$\tilde{\mathbf{y}}_1$	\mathbf{e}_1
\mathbf{x}_i	$\tilde{\mathbf{x}}_{1,i}$	$\mathbf{r}_{1,i}$
\mathbf{X}		\mathbf{X}_{r_1}

Tableau 2 : Opérations dans ERP1 et obtention de ERP2

	Projection sur Δ_2	Projection \perp à Δ_2
	\mathbf{H}_2	$(\mathbf{I} - \mathbf{H}_2)$
\mathbf{e}_1	$\tilde{\mathbf{e}}_2$	$\hat{\mathbf{e}}$

Tableau 3 : Opérations dans ERP2

5.2 TROIS COEFFICIENTS

Dans le cas de 3 coefficients, il y a trois vecteurs-niveaux, x_1 , x_2 et x_3 , portés par 3 droites D_1 , D_2 et D_3 . Ces 3 droites définissent l'ERP initial ou ERP1. La régression classique de y conduit au vecteur \hat{y} situé dans l'ERP1.

On projette y sur une première direction Δ_1 . On obtient un premier vecteur des réponses calculées, $\overrightarrow{OB} = \tilde{y}_1$ (Figure 15). Si l'on choisit une autre droite de projection, on a un second vecteur des réponses calculées. Le lieu des extrémités de ces vecteurs est une sphère S_1 de diamètre OH.

Les vecteurs orthogonaux à OB en B sont situés dans un plan orthogonal à OB en B. Ce plan passe par la droite HA et il coupe la sphère S_1 selon un cercle de diamètre BH. On peut alors décomposer le vecteur \overrightarrow{BH} en 2 vecteurs orthogonaux, \overrightarrow{BC} et \overrightarrow{CH} . On a donc

$$\hat{y} = \overrightarrow{OB} + \overrightarrow{BC} + \overrightarrow{CH}$$

Les 3 vecteurs \overrightarrow{OB} , \overrightarrow{BC} et \overrightarrow{CH} étant orthogonaux entre eux, le vecteur \hat{y} est décomposé en trois vecteurs orthogonaux.

Finalement, pour retrouver \hat{y} en effectuant une régression séquentielle orthogonale, on exécute trois projections successives sur trois directions orthogonales : Δ_1 , Δ_2 et Δ_3 .

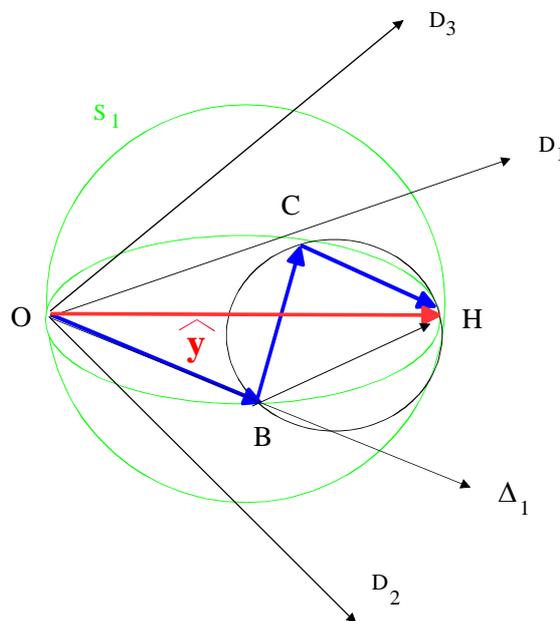


Figure 15 : le vecteur \tilde{y} est décomposé en trois vecteurs orthogonaux.

Projections sur Δ_1

Il y a deux projections (ou régressions) sur Δ_1 de paramètres directeurs t_1 . La première correspond à la régression du vecteur y qui fournit une première solution

\tilde{y}_1 . La seconde correspond à la construction d'un nouvel ERP orthogonal à la droite Δ_1 .

Régression de y sur Δ_1

On choisit la première direction de projection Δ_1 . On obtient, comme précédemment, le vecteur régressé \tilde{y}_1 et le vecteur des résidus e_1 correspondant. On a

$$\tilde{y}_1 = H_1 y$$

$$e_1 = (I - H_1) y$$

avec la matrice de projection

$$H_1 = X t_1 [t_1' X X t_1]^{-1} t_1' X$$

Régression des vecteurs niveaux sur Δ_1

Cette régression a pour objectif d'obtenir un nouvel ERP, l'ERP2, orthogonal à Δ_1 . Cet espace a une dimension de moins que l'ERP1, donc ici deux dimensions. L'ERP2 est donc un plan orthogonal à Δ_1 passant par le point B.

On régresse les trois vecteurs niveaux x_1, x_2 et x_3 sur Δ_1 . On obtient les vecteurs régressés $\tilde{x}_{1,1}, \tilde{x}_{1,2}$ et $\tilde{x}_{1,3}$ qui sont portés par la droite Δ_1 et les vecteurs des résidus $r_{1,1}, r_{1,2}$ et $r_{1,3}$ qui sont orthogonaux à cette même droite

$$\begin{array}{ll} \tilde{x}_{1,1} = H_1 x_1 & r_{1,1} = (I - H_1) x_1 \\ \tilde{x}_{1,2} = H_1 x_2 & r_{1,2} = (I - H_1) x_2 \\ \tilde{x}_{1,3} = H_1 x_3 & r_{1,3} = (I - H_1) x_3 \end{array}$$

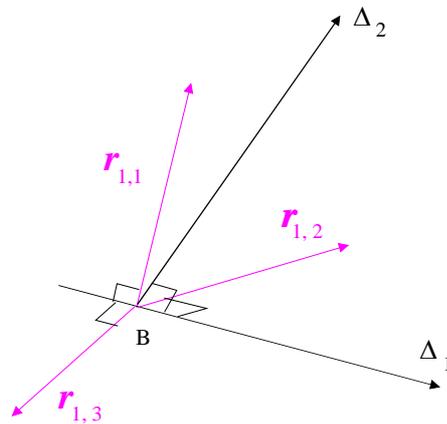


Figure 16 : Les directions $r_{1,i}$ définissent un plan orthogonal à Δ_1 .

Le nouvel espace des réponses possibles, orthogonal à Δ_1 en B, est engendré par les vecteurs des résidus $r_{1,1}, r_{1,2}$ et $r_{1,3}$. On peut les regrouper dans la matrice X_{r_1} telle que

$$\mathbf{X}_{r_1} = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} \end{bmatrix}$$

Les vecteurs des résidus $r_{1,1}$, $r_{1,2}$ et $r_{1,3}$ sont dans le plan orthogonal à la droite Δ_1 et ils définissent l'ERP2 (Figure 16). Ils ne sont pas linéairement indépendants.

On choisit librement les paramètres directeurs de \mathbf{t}_2 qui toutefois ne doivent pas être tous nuls et qui doivent être différents de ceux de \mathbf{t}_1 . On définit ainsi la droite Δ_2 située dans ERP2.

On a la relation

$$\mathbf{X}_{r_1} = (\mathbf{I} - \mathbf{H}_1) \mathbf{X}$$

Projections sur Δ_2

Deux projections (ou régressions) sur Δ_2 vont être considérées. La première projection correspond à la décomposition du vecteur des résidus \mathbf{e}_1 et elle conduit à la seconde solution $\tilde{\mathbf{y}}_2$ de la régression. La seconde projection correspond à la construction d'un nouvel ERP orthogonal à la droite Δ_2 .

Régression de \mathbf{e}_1 sur la direction Δ_2

La régression du premier résidu \mathbf{e}_1 sur la direction Δ_2 conduit au vecteur régressé $\tilde{\mathbf{e}}_2$ porté par Δ_2 et au vecteur des résidus \mathbf{e}_2 , orthogonal à Δ_2 (Figure 17).

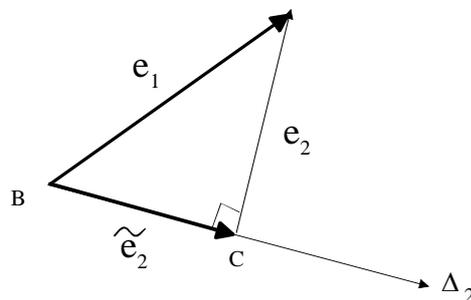


Figure 17 : Le résidu \mathbf{e}_1 est décomposé en deux vecteurs orthogonaux \mathbf{e}_2 et $\tilde{\mathbf{e}}_2$.

Le vecteur $\tilde{\mathbf{e}}_2$, régressé de \mathbf{e}_1 sur Δ_2 , est égal à :

$$\tilde{\mathbf{e}}_2 = \mathbf{H}_2 \mathbf{e}_1 \quad \{20\}$$

Le vecteur des résidus \mathbf{e}_2 correspondant à la régression de \mathbf{e}_1 sur Δ_2 , est égal à :

$$\mathbf{e}_2 = (\mathbf{I} - \mathbf{H}_2) \mathbf{e}_1 \quad \{21\}$$

avec la matrice de projection

$$\mathbf{H}_2 = \mathbf{X}_{r_1} \mathbf{t}_2 \left[\mathbf{t}_2' \mathbf{X}_{r_1} \mathbf{X}_{r_1} \mathbf{t}_2 \right]^{-1} \mathbf{t}_2' \mathbf{X}_{r_1} \quad \{22\}$$

Le premier vecteur des réponses calculées est le vecteur $\tilde{\mathbf{y}}_1$, projection de \mathbf{y} sur la droite Δ_1 . La seconde régression introduit le vecteur $\tilde{\mathbf{e}}_2$ porté par la droite Δ_2 . La somme de ces deux vecteurs est un nouveau vecteur des réponses calculées que l'on peut dénommer $\tilde{\mathbf{y}}_2$ (Figure 18). Ce nouveau vecteur peut être considéré comme une solution de la régression

$$\tilde{\mathbf{y}}_2 = \tilde{\mathbf{y}}_1 + \tilde{\mathbf{e}}_2 = \tilde{\mathbf{y}}_1 + \mathbf{H}_2 \mathbf{e}_1 \quad \{23\}$$

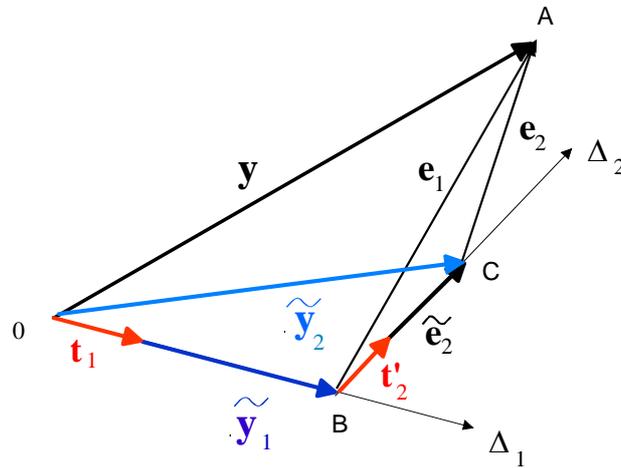


Figure 18 : La deuxième solution $\tilde{\mathbf{y}}_2$ est la somme vectorielle de $\tilde{\mathbf{y}}_1$ et de $\tilde{\mathbf{e}}_2$.

Régression des vecteurs des résidus $\mathbf{r}_{1,i}$ sur Δ_2

Cette régression a pour objectif de définir un nouvel ERP, l'ERP3, orthogonal à Δ_2 . Cet espace possède une dimension de moins que l'espace précédent, donc ici, il n'a qu'une seule dimension. L'ERP3 se réduit à la droite Δ_3 . C'est donc une droite orthogonale à Δ_2 et l'on sait qu'elle passe par le point H de la régression classique.

On régresse les trois vecteurs résidus $\mathbf{r}_{1,1}$, $\mathbf{r}_{1,2}$ et $\mathbf{r}_{1,3}$ sur Δ_2 . On obtient les vecteurs régressés $\tilde{\mathbf{x}}_{2,1}$, $\tilde{\mathbf{x}}_{2,2}$ et $\tilde{\mathbf{x}}_{2,3}$ qui sont portés par la droite Δ_2 et de nouveaux vecteurs des résidus $\mathbf{r}_{2,1}$, $\mathbf{r}_{2,2}$ et $\mathbf{r}_{2,3}$ qui sont orthogonaux à cette même droite

$$\begin{aligned} \tilde{\mathbf{x}}_{2,1} &= \mathbf{H}_2 \mathbf{r}_{1,1} & \mathbf{r}_{2,1} &= (\mathbf{I} - \mathbf{H}_2) \mathbf{r}_{1,1} \\ \tilde{\mathbf{x}}_{2,2} &= \mathbf{H}_2 \mathbf{r}_{1,2} & \mathbf{r}_{2,2} &= (\mathbf{I} - \mathbf{H}_2) \mathbf{r}_{1,2} \\ \tilde{\mathbf{x}}_{2,3} &= \mathbf{H}_2 \mathbf{r}_{1,3} & \mathbf{r}_{2,3} &= (\mathbf{I} - \mathbf{H}_2) \mathbf{r}_{1,3} \end{aligned}$$

Le nouvel espace des réponses possibles est orthogonal à Δ_2 en C et il est engendré par les vecteurs des résidus $\mathbf{r}_{2,1}$, $\mathbf{r}_{2,2}$ et $\mathbf{r}_{2,3}$. Ces vecteurs sont regroupés dans la matrice \mathbf{X}_{r_2} telle que

$$\mathbf{X}_{r_2} = \begin{bmatrix} \mathbf{r}_{2,1} & \mathbf{r}_{2,2} & \mathbf{r}_{2,3} \end{bmatrix}$$

Les vecteurs des résidus $r_{2,1}$, $r_{2,2}$ et $r_{2,3}$ sont orthogonaux à la droite Δ_2 . Ils sont parallèles entre eux et définissent l'ERP3 qui ne possède qu'une seule dimension. Les vecteurs des résidus $r_{2,1}$, $r_{2,2}$ et $r_{2,3}$ ne sont pas linéairement indépendants et la matrice X_{r_2} n'est pas de plein rang.

On a

$$X_{r_2} = (I - H_2) X_{r_1} \quad \{24\}$$

$$X_{r_2} = (I - H_2) (I - H_1) X \quad \{25\}$$

Projections sur Δ_3

C'est la dernière projection (ou régression) et elle sera exécutée sur Δ_3 . Elle correspond à la décomposition du vecteur des résidus e_2 et elle conduit à la solution classique de la régression : \hat{y} .

Régression de e_2 sur la direction Δ_3

On obtient le vecteur régressé (Figure 19):

$$\overrightarrow{CH} = \tilde{e}_3 = H_3 e_2 \quad \{26\}$$

Le vecteur des résidus e_3 correspondant à la régression de e_2 sur Δ_3 , est égal à :

$$e_3 = (I - H_3) e_2 \quad \{27\}$$

avec

$$H_3 = X_{r_2} t_3 [t_3' X_{r_2} X_{r_2} t_3]^{-1} t_3' X_{r_2} \quad \{28\}$$

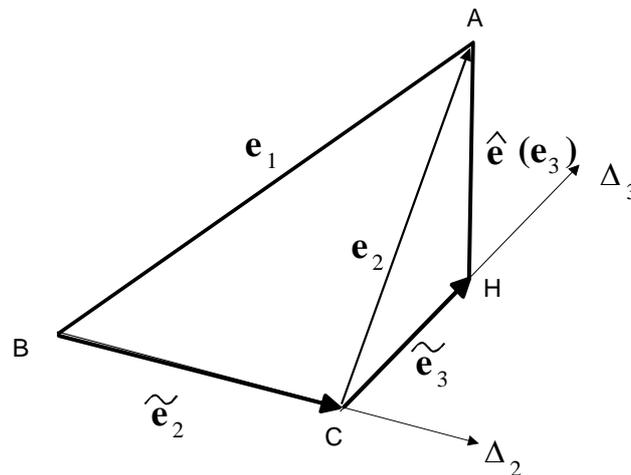


Figure 19 : Le vecteur e_2 est décomposé en un vecteur régressé $\overrightarrow{CH} = \tilde{e}_3$ et un vecteur des résidus, e_3 , qui est ici \hat{e} .

Réduction dimensionnelle des ERP

La projection $(\mathbf{I}-\mathbf{H}_1)$ projette les vecteurs de l'ERP1 sur un plan orthogonal à Δ_1 . Elle transforme le vecteur \mathbf{t}_2 défini dans ERP1, en un vecteur \mathbf{t}'_2 orthogonal à Δ_1 et situé dans l'ERP2 qui ne possède que deux dimensions. Pour faire les calculs on passe dans le référentiel des réponses, le vecteur \mathbf{t}_2 est transformé en $\mathbf{X t}_2$. On a

$$\mathbf{X t}'_2 = (\mathbf{I}-\mathbf{H}_1) \mathbf{X t}_2$$

La projection $(\mathbf{I}-\mathbf{H}_1)$ transforme le vecteur \mathbf{t}_3 de ERP1 (trois dimensions) en un vecteur \mathbf{t}'_3 de ERP2 (deux dimensions).

$$\mathbf{X t}'_3 = (\mathbf{I}-\mathbf{H}_1) \mathbf{X t}_3$$

La projection $(\mathbf{I}-\mathbf{H}_2)$ transforme le vecteur \mathbf{t}'_3 de ERP2 (deux dimensions) en un vecteur \mathbf{t}''_3 de ERP3 (une dimension).

$$\mathbf{X t}''_3 = (\mathbf{I}-\mathbf{H}_2) \mathbf{X t}'_3 = (\mathbf{I}-\mathbf{H}_2)(\mathbf{I}-\mathbf{H}_1) \mathbf{X t}_3$$

On peut donc choisir au mieux des intérêts de l'étude les directions Δ_i définies par les paramètres directeurs \mathbf{t}_i .

Les solutions de la régression

Il y a trois vecteurs candidats comme solution de la régression : le vecteur $\tilde{\mathbf{y}}_1$ projeté sur Δ_1 , le vecteur $\tilde{\mathbf{y}}_2$ qui résulte de la seconde projection et le vecteur $\hat{\mathbf{y}}$ qui est la solution classique. On a

$$\tilde{\mathbf{y}}_1 = \mathbf{H}_1 \mathbf{y}$$

$$\tilde{\mathbf{y}}_2 = \tilde{\mathbf{y}}_1 + \tilde{\mathbf{e}}_2 = \mathbf{H}_1 \mathbf{y} + \mathbf{H}_2 \mathbf{e}_1$$

$$\hat{\mathbf{y}} = \tilde{\mathbf{y}}_2 + \tilde{\mathbf{e}}_3 = \tilde{\mathbf{y}}_1 + \tilde{\mathbf{e}}_2 + \tilde{\mathbf{e}}_3$$

$$\hat{\mathbf{y}} = \mathbf{H}_1 \mathbf{y} + \mathbf{H}_2 \mathbf{e}_1 + \mathbf{H}_3 \mathbf{e}_2$$

Pour faciliter la compréhension, on peut représenter les trois vecteurs solutions dans un plan ne prenant en compte que les vecteurs $\tilde{\mathbf{y}}_i$ et les vecteurs $\tilde{\mathbf{e}}_i$ (Figure 20).

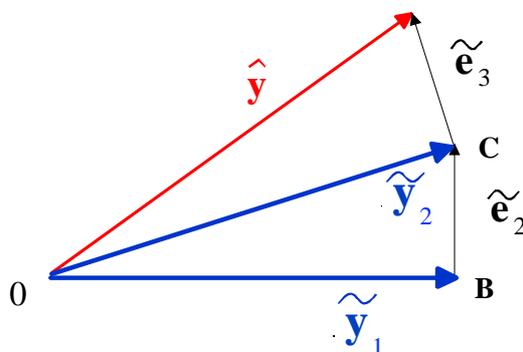


Figure 20 : Illustration, dans un plan, des trois solutions de la régression.

Coefficients des vecteurs régressés

Les coefficients $\tilde{\mathbf{b}}_1$ du modèle mathématique donnant le premier vecteur régressé $\tilde{\mathbf{y}}_1$, sont fournis par la relation

$$\tilde{\mathbf{b}}_1 = \tilde{k} \mathbf{t}_1 = \mathbf{t}_1 \left[{}^t \mathbf{t}_1 \mathbf{X} \mathbf{X} \mathbf{t}_1 \right]^{-1} {}^t \mathbf{t}_1 \mathbf{X} \mathbf{y}$$

Les coefficients $\tilde{\mathbf{b}}_2$ du modèle mathématique donnant le second vecteur régressé $\tilde{\mathbf{y}}_2$, sont fournis par la relation :

$$\tilde{\mathbf{b}}_2 = \left[{}^t \mathbf{X} \mathbf{X} \right]^{-1} {}^t \mathbf{X} \tilde{\mathbf{y}}_2$$

Les coefficients $\hat{\mathbf{a}}$ du modèle mathématique donnant le troisième vecteur régressé $\hat{\mathbf{y}}$, sont fournis par la relation classique :

$$\hat{\mathbf{a}} = \left[{}^t \mathbf{X} \mathbf{X} \right]^{-1} {}^t \mathbf{X} \mathbf{y}$$

Résumé des principales opérations

Le

Tableau 4, le Tableau 5 et le Tableau 6 indiquent les principales opérations effectuées dans ERP1, ERP2 et ERP3.

	Projection sur Δ_1	Projection \perp à Δ_1
	\mathbf{H}_1	$(\mathbf{I} - \mathbf{H}_1)$
\mathbf{y}	$\tilde{\mathbf{y}}_1$	\mathbf{e}_1
\mathbf{x}_i	$\tilde{\mathbf{x}}_{1,i}$	$\mathbf{r}_{1,i}$
\mathbf{X}		\mathbf{X}_{r_1}

Tableau 4 : Opérations dans ERP1 et obtention de ERP2

	Projection sur Δ_2	Projection \perp à Δ_2
	\mathbf{H}_2	$(\mathbf{I} - \mathbf{H}_2)$
\mathbf{e}_1	$\tilde{\mathbf{e}}_2$	\mathbf{e}_2
$\mathbf{r}_{1,i}$	$\tilde{\mathbf{x}}_{2,i}$	$\mathbf{r}_{2,i}$
\mathbf{X}_{r_1}		\mathbf{X}_{r_2}

Tableau 5 : Opérations dans ERP2 et obtention de ERP3

	Projection sur Δ_3	Projection \perp à Δ_3
	\mathbf{H}_3	$(\mathbf{I} - \mathbf{H}_3)$
\mathbf{e}_2	$\tilde{\mathbf{e}}_3$	$\hat{\mathbf{e}}$

Tableau 6 : Opérations dans ERP3

5.3 N COEFFICIENTS

Les raisonnements que nous avons tenus peuvent s'appliquer lorsqu'il y a n coefficients. Dans ce cas, on a n vecteurs-niveaux, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, portés par n droites D_1, D_2, \dots, D_n . Ces n droites définissent l'ERP1. La régression classique de \mathbf{y} conduit au vecteur $\hat{\mathbf{y}}$ situé dans l'ERP1.

On projette \mathbf{y} sur une première direction Δ_1 . On obtient un premier vecteur des réponses calculées \overrightarrow{OB} . Si l'on considère toutes les directions possibles, le lieu des extrémités des vecteurs $\overrightarrow{OB_i}$ est une hypersphère S_1 à p dimensions de diamètre OH . Les vecteurs orthogonaux à OB en B sont dans un hyperplan orthogonal à OB en B . Cet hyperplan passe par la droite HA et il coupe l'hypersphère S_1 selon une hypersphère S_2 de dimensions $p-1$ de diamètre BH . On peut alors décomposer BH en $p-1$ vecteurs orthogonaux en passant par des hypersphères dont les dimensions diminuent d'une unité à chaque fois. On termine ce processus de décomposition par un cercle et deux vecteurs orthogonaux. Le vecteur $\hat{\mathbf{y}}$ peut ainsi être décomposé en p vecteurs orthogonaux.

Dans l'ERP1, on régresse \mathbf{y} sur une direction Δ_1 . On obtient un vecteur régressé, $\tilde{\mathbf{y}}_1$ porté par Δ_1 , et un vecteur résidu \mathbf{e}_1 situé dans l'hyperplan orthogonal à Δ_1 en B . On a

$$\tilde{\mathbf{y}}_1 = \mathbf{H}_1 \mathbf{y}$$

$$\mathbf{e}_1 = (\mathbf{I} - \mathbf{H}_1) \mathbf{y}$$

On définit le nouvel ERP, l'ERP2, en régressant les vecteurs niveaux $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ sur Δ_1 . Les vecteurs des résidus obtenus $\mathbf{r}_{1,1}, \mathbf{r}_{1,2}$ et $\mathbf{r}_{1,p}$ engendrent l'ERP2 qui à $p-1$ dimensions et qui est orthogonal à Δ_1 . On définit une direction de projection Δ_2 sur laquelle on projette le résidu \mathbf{e}_1 . On obtient un vecteur régressé $\tilde{\mathbf{e}}_2$ porté par la droite Δ_2 et un vecteur des résidus \mathbf{e}_2 orthogonal à Δ_2 .

$$\tilde{\mathbf{e}}_2 = \mathbf{H}_2 \mathbf{e}_1$$

$$\mathbf{e}_2 = (\mathbf{I} - \mathbf{H}_2) \mathbf{e}_1$$

On obtient ainsi une seconde solution de la régression

$$\tilde{\mathbf{y}}_2 = \tilde{\mathbf{y}}_1 + \tilde{\mathbf{e}}_2 = \tilde{\mathbf{y}}_1 + \mathbf{H}_2 \mathbf{e}_1$$

On peut poursuivre la décomposition en définissant un troisième ERP, l'ERP3. On obtient ERP3 en régressant les vecteurs des résidus $\mathbf{r}_{1,1}, \mathbf{r}_{1,2}$ et $\mathbf{r}_{1,p}$ sur Δ_2 . Cette

régression fournit les résidus $r_{2,1}$, $r_{2,2}$ et $r_{2,p}$ qui sont orthogonaux à Δ_2 et qui engendrent l'ERP3 qui ne comprend plus que $p-2$ dimensions. On choisit une troisième direction de projection Δ_3 .

Le vecteur des résidus e_2 est projeté sur la nouvelle direction, Δ_3 , orthogonale à Δ_1 et à Δ_2 . On obtient un vecteur projeté, \tilde{e}_3 , et un nouveau vecteur des résidus, e_3 (Figure 21).

$$\tilde{e}_3 = \mathbf{H}_3 e_2$$

$$e_3 = (\mathbf{I} - \mathbf{H}_3) e_2$$

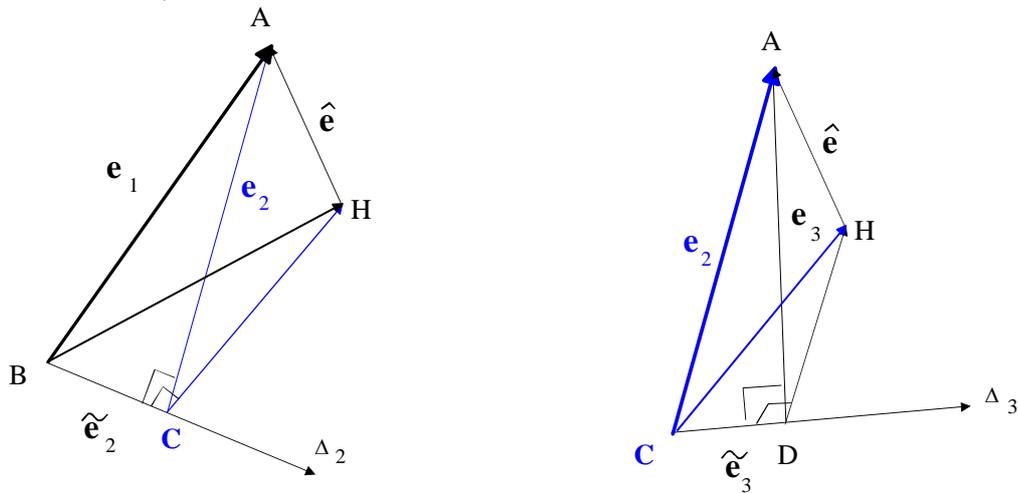


Figure 21 : A gauche le vecteur e_1 est régressé sur la direction Δ_2 . Le vecteur des résidus e_2 de cette régression est à son tour régressé sur direction Δ_3 (figure de droite). Le vecteur des résidus e_3 ainsi obtenu sera, lui aussi, régressé sur une nouvelle direction Δ_4 .

On obtient ainsi une troisième solution, \tilde{y}_3 (Figure 22) :

$$\tilde{y}_3 = \tilde{y}_2 + \tilde{e}_3 = \tilde{y}_1 + \tilde{e}_2 + \tilde{e}_3$$

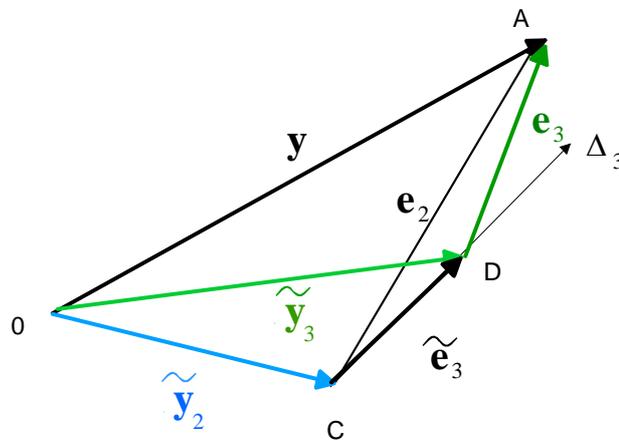


Figure 22 : La troisième solution \tilde{y}_3 est la somme vectorielle de \tilde{y}_2 et de \tilde{e}_3 .

A son tour le résidu, \mathbf{e}_3 est décomposé en un vecteur projeté, $\tilde{\mathbf{e}}_4$, et un vecteur des résidus, \mathbf{e}_4 . De décomposition en décomposition, on arrive à retrouver le vecteur $\hat{\mathbf{y}}$ de la régression classique. La Figure 23 illustre les solutions successives proposées par la régression séquentielle orthogonale.

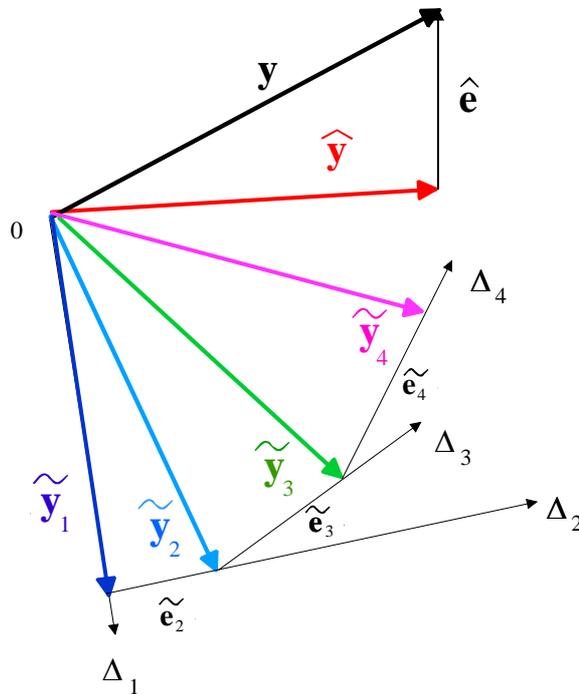


Figure 23 : La régression séquentielle orthogonale propose plusieurs solutions pour lier la réponse aux facteurs.

Réduction dimensionnelle des ERP

La projection $(\mathbf{I} - \mathbf{H}_1)$ projette les vecteurs de ERP1 sur un espace orthogonal à Δ_1 . Elle transforme le vecteur \mathbf{t}_i (p composantes dans ERP1 et n composantes dans le référentiel des réponses) de l'espace ERP1 de p dimensions en un vecteur \mathbf{t}'_i orthogonal à Δ_1 et situé dans l'espace ERP2 de $p-1$ dimensions. Elle transforme la matrice \mathbf{X} en une matrice \mathbf{X}_{r_1} dont les colonnes, considérées comme des vecteurs, définissent ERP2.

$$\mathbf{X} \mathbf{t}'_i = (\mathbf{I} - \mathbf{H}_1) \mathbf{X} \mathbf{t}_i \quad \mathbf{X}_{r_1} = (\mathbf{I} - \mathbf{H}_1) \mathbf{X}$$

La projection $(\mathbf{I} - \mathbf{H}_2)$ transforme le vecteur \mathbf{t}'_i de ERP2 ($p-1$ dimensions) en un vecteur \mathbf{t}''_i de ERP3 ($p-2$ dimensions). Elle transforme \mathbf{X}_{r_1} en \mathbf{X}_{r_2} qui définit ERP3.

$$\begin{aligned} \mathbf{X} \mathbf{t}''_i &= (\mathbf{I} - \mathbf{H}_2) \mathbf{X} \mathbf{t}'_i & \text{ou} & & \mathbf{X} \mathbf{t}''_i &= (\mathbf{I} - \mathbf{H}_2)(\mathbf{I} - \mathbf{H}_1) \mathbf{X} \mathbf{t}_i \\ \mathbf{X}_{r_2} &= (\mathbf{I} - \mathbf{H}_2) \mathbf{X}_{r_1} & \text{ou} & & \mathbf{X}_{r_2} &= (\mathbf{I} - \mathbf{H}_2) (\mathbf{I} - \mathbf{H}_1) \mathbf{X} \end{aligned}$$

Les projections se poursuivent jusqu'à obtenir l'ERP(j). Le vecteur $\mathbf{t}_i^{(j-1)'}$ est projeté ($j-1$) fois, on a donc

$$\mathbf{X} \mathbf{t}_i^{(j-1)'} = (\mathbf{I} - \mathbf{H}_{j-1}) \cdots (\mathbf{I} - \mathbf{H}_2)(\mathbf{I} - \mathbf{H}_1) \mathbf{X} \mathbf{t}_i$$

La matrice $\mathbf{X}_{r_{j-1}}$ est obtenue après $(j-1)$ projections, on a donc

$$\mathbf{X}_{r_{j-1}} = (\mathbf{I} - \mathbf{H}_{j-1}) \cdots (\mathbf{I} - \mathbf{H}_2) (\mathbf{I} - \mathbf{H}_1) \mathbf{X}$$

Les solutions de la régression

Il y a p solutions de la régression depuis le vecteur $\tilde{\mathbf{y}}_1$ projeté sur Δ_1 jusqu'au vecteur $\hat{\mathbf{y}}$ de la solution classique. On a simplement

$$\tilde{\mathbf{y}}_k = \tilde{\mathbf{y}}_1 + \sum_{i=2}^{i=k} \tilde{\mathbf{e}}_i$$

On peut également écrire les solutions successives avec les matrices de projections :

$$\tilde{\mathbf{y}}_1 = \mathbf{H}_1 \mathbf{y}$$

$$\tilde{\mathbf{y}}_2 = \tilde{\mathbf{y}}_1 + \tilde{\mathbf{e}}_2 = \tilde{\mathbf{y}}_1 + \mathbf{H}_2 \mathbf{e}_1 = \mathbf{H}_1 \mathbf{y} + \mathbf{H}_2(\mathbf{I} - \mathbf{H}_1) \mathbf{y}$$

$$\tilde{\mathbf{y}}_3 = \mathbf{H}_1 \mathbf{y} + \mathbf{H}_2(\mathbf{I} - \mathbf{H}_1) \mathbf{y} + \mathbf{H}_3(\mathbf{I} - \mathbf{H}_2) (\mathbf{I} - \mathbf{H}_1) \mathbf{y}$$

etc.

La régression séquentielle orthogonale propose donc plusieurs solutions possibles puisque chaque vecteur $\tilde{\mathbf{y}}_i$ représente un modèle mathématique permettant d'obtenir des réponses calculées proches des réponses mesurées.

On peut représenter, dans un plan, toutes les solutions de la régression en ne prenant en compte que les vecteurs $\tilde{\mathbf{y}}_i$ et les vecteurs $\tilde{\mathbf{e}}_i$ (Figure 24).

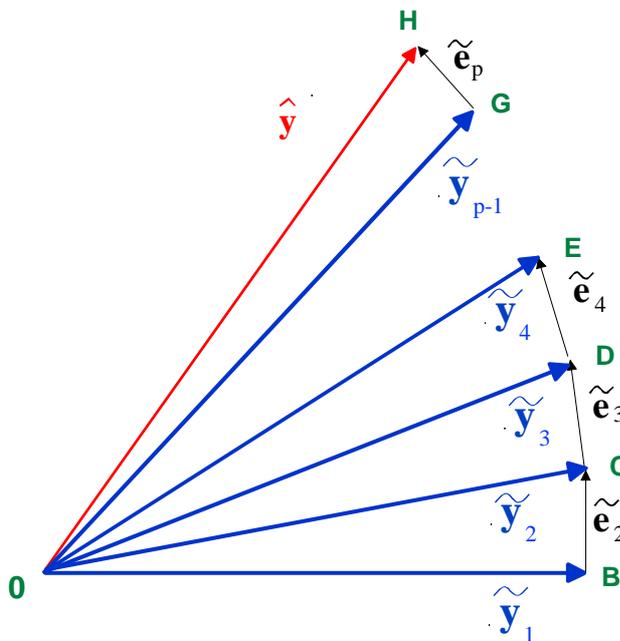


Figure 24 : Illustration, dans un plan, des p solutions de la régression

Coefficients des vecteurs régressés

Les coefficients $\hat{\mathbf{a}}$ du modèle mathématique donnant le vecteur régressé classique $\hat{\mathbf{y}}$, sont fournis par la relation classique :

$$\hat{\mathbf{a}} = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{y}$$

Les coefficients $\tilde{\mathbf{b}}_i$ du modèle mathématique correspondant aux diverses solutions possibles sont donnés par

$$\tilde{\mathbf{b}}_i = [\mathbf{X}^t \mathbf{X}]^{-1} \mathbf{X}^t \tilde{\mathbf{y}}_i$$

6. REGRESSION PLS1

La régression PLS1 est une simple application de la régression séquentielle orthogonale. Dans ce cas, on décide que les paramètres directeurs de la première direction de projection Δ_1 sont proportionnels aux covariances entre le vecteur des réponses mesurées, \mathbf{y} , et les vecteurs-niveaux. On a donc

$$\mathbf{t}_1 = \begin{bmatrix} t_{1,1} \\ t_{1,2} \\ \vdots \\ t_{1,p} \end{bmatrix} = \kappa_1 \begin{bmatrix} Cov(\mathbf{x}_1, \mathbf{y}) \\ Cov(\mathbf{x}_2, \mathbf{y}) \\ \vdots \\ Cov(\mathbf{x}_p, \mathbf{y}) \end{bmatrix} = \kappa_1 \mathbf{Cov}_1 \quad \{29\}$$

- \mathbf{t}_1 matrice $(p,1)$ des paramètres directeurs de la droite Δ_1 sur laquelle on projette.
- κ_1 scalaire de proportionnalité.
- \mathbf{Cov}_1 matrice $(p,1)$ des covariances entre le vecteur des réponses mesurées, \mathbf{y} , et les vecteurs-niveaux, \mathbf{x}_i . Ces covariances sont connues.

Dans ce cas la matrice de projection est égale à

$$\mathbf{H}_1 = \mathbf{X} \mathbf{Cov}_1 \left[\mathbf{Cov}_1^t \mathbf{X}^t \mathbf{X} \mathbf{Cov}_1 \right]^{-1} \mathbf{Cov}_1^t \mathbf{X}$$

D'où le vecteur $\tilde{\mathbf{y}}_1$, projection de \mathbf{y} sur la droite Δ_1 :

$$\tilde{\mathbf{y}}_1 = \mathbf{X} \mathbf{Cov}_1 \left[\mathbf{Cov}_1^t \mathbf{X}^t \mathbf{X} \mathbf{Cov}_1 \right]^{-1} \mathbf{Cov}_1^t \mathbf{X} \mathbf{y}$$

Ce vecteur des réponses calculées, $\tilde{\mathbf{y}}_1$, obtenu par régression sur la droite Δ_1 , est appelé premier vecteur des réponses PLS 1.

Le vecteur des résidus \mathbf{e}_1 est situé dans l'hyperplan orthogonal à Δ_1 en B et il est égal à

$$\mathbf{e}_1 = (\mathbf{I} - \mathbf{H}_1) \mathbf{y}$$

C'est ce vecteur des résidus, \mathbf{e}_1 , qui va être décomposé à nouveau dans un ERP de $p-1$ dimensions défini par les premiers vecteurs des résidus $\mathbf{r}_{1,1}, \mathbf{r}_{1,2}, \dots, \mathbf{r}_{1,p}$. Dans la régression PLS1, on définit la seconde direction de projection par les covariances entre le vecteur des résidus \mathbf{e}_1 et les vecteurs des résidus $\mathbf{r}_{1,i}$.

$$\mathbf{t}_2 = \begin{bmatrix} t_{2,1} \\ t_{2,2} \\ \vdots \\ t_{2,p} \end{bmatrix} = \begin{bmatrix} \text{Cov}(\mathbf{r}_{1,1}, \mathbf{e}_1) \\ \text{Cov}(\mathbf{r}_{1,2}, \mathbf{e}_1) \\ \vdots \\ \text{Cov}(\mathbf{r}_{1,p}, \mathbf{e}_1) \end{bmatrix} = \kappa_2 \mathbf{Cov}_2$$

On obtient un vecteur régressé $\tilde{\mathbf{e}}_2$ et un vecteur des résidus \mathbf{e}_2 . Comme dans la régression séquentielle orthogonale, on poursuit les décompositions jusqu'à obtenir le vecteur $\hat{\mathbf{y}}$ de la régression classique. Pour faciliter le choix de la meilleure solution il existe des critères statistiques tels que le Q^2 ou le PRESS. L'idée générale de ces critères est de s'arrêter quand les vecteurs $\tilde{\mathbf{e}}_i$ que l'on ajoute deviennent égaux ou inférieurs à l'erreur.

On retrouve toutes les relations que nous avons indiquées en remplaçant \mathbf{t}_i par sa valeur :

$$\mathbf{t}_i = \begin{bmatrix} t_{i,1} \\ t_{i,2} \\ \vdots \\ t_{i,p} \end{bmatrix} = \begin{bmatrix} \text{Cov}(\mathbf{r}_{i-1,1}, \mathbf{e}_{i-1}) \\ \text{Cov}(\mathbf{r}_{i-1,2}, \mathbf{e}_{i-1}) \\ \vdots \\ \text{Cov}(\mathbf{r}_{i-1,p}, \mathbf{e}_{i-1}) \end{bmatrix} = \kappa_i \mathbf{Cov}_i \quad \{30\}$$

Si la régression PLS conduit à ne retenir que deux axes de projections pour obtenir le modèle le plus proche de la meilleure solution, il est possible avec la RLSO de n'effectuer qu'une seule projection. En effet, il suffit de projeter le vecteur des réponses mesurées sur un axe dont les paramètres directeurs sont proportionnels aux coefficients du vecteur solution de la PLS. Dans ce cas on pourra choisir d'autres directions de projection de manière à entourer la solutions PLS. L'estimation du meilleur modèle sera ainsi réalisée sur des projections voisines de la solution PLS mais pouvant être meilleures encore. La régression PLS ne donne que des solutions discontinues alors que la régression RLSO permet d'obtenir toutes les solutions possibles et de choisir la meilleure en accord avec le critère d'optimisation.

7. COMPARAISON DES REGRESSIONS PLS1 ET RLSO

La différence la plus importante entre ces deux régressions séquentielles est le libre choix de la direction de projection offert par la RLSO. En effet, les deux régressions offrent le même nombre de solutions possibles mais la PLS1 impose l'emplacement de ces solutions dans l'espace défini par le référentiel des réponses. Au contraire, la RLSO offre la possibilité de placer les solutions en une infinité de positions. Il est ainsi possible de regrouper les solutions dans un faible espace favorable et de choisir la meilleure solution. Ce choix peut se faire en utilisant les critères mis au point pour la PLS1 et en les associant à des plans d'expériences pour trouver la meilleure solution ou le meilleur compromis possible. Avec la RLSO, il est aussi possible d'imposer des contraintes comme, par exemple, la valeur d'un ou de plusieurs coefficients. En particulier il est facile d'éliminer tous les coefficients non significatifs et de trouver la meilleure solution utilisant l'influence de quelques facteurs et répondant à des impératifs bien précis. La souplesse de la RLSO s'oppose à la rigidité de la PLS1 et offre par là une multitude de possibilités nouvelles à tous les utilisateurs de la régression.

ANNEXE 1

CALCUL DU COEFFICIENT \tilde{k} DE LA REGRESSION SUR UNE DIRECTION

La solution des moindres carrés est la suivante :

$${}^t \mathbf{e} \mathbf{e} = {}^t (\mathbf{y} - \mathbf{X} \mathbf{b}) (\mathbf{y} - \mathbf{X} \mathbf{b})$$

$${}^t \mathbf{e} \mathbf{e} = ({}^t \mathbf{y} - {}^t \mathbf{b}' \mathbf{X}) (\mathbf{y} - \mathbf{X} \mathbf{b})$$

$${}^t \mathbf{e} \mathbf{e} = {}^t \mathbf{y} \mathbf{y} - {}^t \mathbf{y} \mathbf{X} \mathbf{b} - {}^t \mathbf{b}' \mathbf{X} \mathbf{y} + {}^t \mathbf{b}' \mathbf{X} \mathbf{X} \mathbf{b}$$

$${}^t \mathbf{e} \mathbf{e} = {}^t \mathbf{y} \mathbf{y} - {}^t \mathbf{y} \mathbf{X} k \mathbf{t}_\Delta - k {}^t \mathbf{t}_\Delta {}^t \mathbf{X} \mathbf{y} + k^2 {}^t \mathbf{t}_\Delta {}^t \mathbf{X} \mathbf{X} \mathbf{t}_\Delta$$

$$\frac{\partial {}^t \mathbf{e} \mathbf{e}}{\partial k} = -{}^t \mathbf{y} \mathbf{X} \mathbf{t}_\Delta - {}^t \mathbf{t}_\Delta {}^t \mathbf{X} \mathbf{y} + 2k {}^t \mathbf{t}_\Delta {}^t \mathbf{X} \mathbf{X} \mathbf{t}_\Delta$$

$$-{}^t \mathbf{y} \mathbf{X} \mathbf{t}_\Delta - {}^t \mathbf{t}_\Delta {}^t \mathbf{X} \mathbf{y} + 2\tilde{k} {}^t \mathbf{t}_\Delta {}^t \mathbf{X} \mathbf{X} \mathbf{t}_\Delta = 0$$

Or, \mathbf{y} étant une matrice (n,1) ; \mathbf{X} une (n,p) et \mathbf{t} une (p,1), on a

$${}^t \mathbf{y} \mathbf{X} \mathbf{t}_\Delta = {}^t \mathbf{t}_\Delta {}^t \mathbf{X} \mathbf{y}$$

car un scalaire est égal à son transposé. D'où

$${}^t \mathbf{t}_\Delta {}^t \mathbf{X} \mathbf{y} = \tilde{k} {}^t \mathbf{t}_\Delta {}^t \mathbf{X} \mathbf{X} \mathbf{t}_\Delta$$

$$\tilde{k} = \left[{}^t \mathbf{t}_\Delta {}^t \mathbf{X} \mathbf{X} \mathbf{t}_\Delta \right]^{-1} {}^t \mathbf{t}_\Delta {}^t \mathbf{X} \mathbf{y}$$

\tilde{k} est bien un scalaire :

$$(1,p)(p,n)(n,p)(p,1)(1,p)(p,n)(n,1) = (1,1)$$

D'où les composantes du vecteur régressé

$$\tilde{\mathbf{b}} = \tilde{k} \mathbf{t}_\Delta$$

ANNEXE 2

DEFINITION DES DIRECTIONS

Les directions ont été définies par les paramètres directeurs des vecteurs unitaires \mathbf{t} . Mais il est tout à fait possible d'utiliser d'autres vecteurs à condition, bien sûr, qu'ils possèdent la même direction que \mathbf{t} . Si on a utilisé un vecteur \mathbf{v} parallèle à \mathbf{t} mais n'étant pas unitaire, on a

$$\mathbf{t} = \kappa \mathbf{v}$$

Les coefficients du modèle mathématique s'écrivent alors

$$\tilde{\mathbf{b}} = \tilde{k}_1 \mathbf{t} = \kappa \mathbf{v} \left[\kappa {}^t \mathbf{v} {}^t \mathbf{X} \mathbf{X} \kappa \mathbf{v} \right]^{-1} \kappa {}^t \mathbf{v} {}^t \mathbf{X} \mathbf{y}$$

soit

$$\tilde{\mathbf{b}} = \mathbf{v} \left[{}^t \mathbf{v} {}^t \mathbf{X} \mathbf{X} \mathbf{v} \right]^{-1} {}^t \mathbf{v} {}^t \mathbf{X} \mathbf{y}$$

Le coefficient de proportionnalité a disparu de la matrice des coefficients.

Considérons maintenant la matrice de projection

$$\mathbf{H} = \mathbf{X} \mathbf{K} \mathbf{v} \left[\mathbf{K} \mathbf{v} \mathbf{v}' \mathbf{X} \mathbf{X} \mathbf{K} \mathbf{v} \right]^{-1} \mathbf{K} \mathbf{v} \mathbf{v}' \mathbf{X}$$

soit

$$\mathbf{H} = \mathbf{X} \mathbf{v} \left[\mathbf{v} \mathbf{v}' \mathbf{X} \mathbf{X} \mathbf{v} \right]^{-1} \mathbf{v} \mathbf{v}' \mathbf{X}$$

Le coefficient de proportionnalité a disparu de la matrice de projection.

A la place de \mathbf{t} , on peut employer les caractéristiques d'un vecteur parallèle.

ANNEXE 3

ORTHOGONALITE DES DIRECTIONS DE PROJECTION

On peut vérifier l'orthogonalité des directions Δ_i . On opère dans le référentiel orthonormé des réponses et on utilise les transformés $\mathbf{X} \mathbf{t}_i$ des vecteurs \mathbf{t}_i . Calculons le produit scalaire PS du vecteur \mathbf{t}_1 de l'ERP1 avec le vecteur \mathbf{t}_2 projeté dans l'ERP2. On a

$$PS = \left[\mathbf{t}_1 \mathbf{v}' \mathbf{X} \mathbf{X} \mathbf{v} \mathbf{t}_2 \right]$$

soit

$$PS = \left[\mathbf{t}_1 \mathbf{v}' \mathbf{X} (\mathbf{I} - \mathbf{H}_1) \mathbf{X} \mathbf{t}_2 \right]$$

$$\mathbf{H}_1 = \mathbf{X} \mathbf{t}_1 \left[\mathbf{t}_1 \mathbf{v}' \mathbf{X} \mathbf{X} \mathbf{t}_1 \right]^{-1} \mathbf{t}_1 \mathbf{v}' \mathbf{X}$$

$$PS = \left[\mathbf{t}_1 \mathbf{v}' \mathbf{X} (\mathbf{I} - \mathbf{X} \mathbf{t}_1 \left[\mathbf{t}_1 \mathbf{v}' \mathbf{X} \mathbf{X} \mathbf{t}_1 \right]^{-1} \mathbf{t}_1 \mathbf{v}' \mathbf{X}) \mathbf{X} \mathbf{t}_2 \right]$$

$$PS = \left[\left(\mathbf{t}_1 \mathbf{v}' \mathbf{X} - \mathbf{t}_1 \mathbf{v}' \mathbf{X} \mathbf{X} \mathbf{t}_1 \left[\mathbf{t}_1 \mathbf{v}' \mathbf{X} \mathbf{X} \mathbf{t}_1 \right]^{-1} \mathbf{t}_1 \mathbf{v}' \mathbf{X} \right) \mathbf{X} \mathbf{t}_2 \right]$$

$$PS = \left[\left(\mathbf{t}_1 \mathbf{v}' \mathbf{X} - \mathbf{I} \mathbf{t}_1 \mathbf{v}' \mathbf{X} \right) \mathbf{X} \mathbf{t}_2 \right]$$

$$PS = \left[\left(\mathbf{t}_1 \mathbf{v}' \mathbf{X} - \mathbf{t}_1 \mathbf{v}' \mathbf{X} \right) \mathbf{X} \mathbf{t}_2 \right]$$

$$PS = \mathbf{0} (\mathbf{X} \mathbf{t}_2) = 0$$

ANNEXE 4

Dans la Figure 24 les triangles OBC, OCD, ... OGH sont tous rectangles. On a donc :

$$\|\tilde{\mathbf{y}}_1\| \leq \|\tilde{\mathbf{y}}_2\| \leq \|\tilde{\mathbf{y}}_3\| \leq \dots \leq \|\tilde{\mathbf{y}}_{p-1}\| \leq \|\hat{\mathbf{y}}\|$$

C'est donc le vecteur de la régression classique qui a la plus grande norme.

ANNEXE 5

Le triangle OBA de la figure 6 est rectangle, on a donc la relation de Pythagore :

$$\mathbf{y} = \tilde{\mathbf{y}}_1 + \mathbf{e}_1$$

$$\mathbf{y}' \mathbf{y} = \tilde{\mathbf{y}}_1' \tilde{\mathbf{y}}_1 + \mathbf{e}_1' \mathbf{e}_1$$

Quand on centre et que l'on réduit les réponses brutes (avec $n-1$ pour le calcul de la variance), la somme des carrés des réponses centrées réduites est égale à $n-1$, soit

$${}^t \mathbf{y} \mathbf{y} = n - 1$$

D'après la figure 17, on a donc la relation de Pythagore :

$${}^t \mathbf{e}_1 \mathbf{e}_1 = {}^t \tilde{\mathbf{e}}_2 \tilde{\mathbf{e}}_2 + {}^t \mathbf{e}_2 \mathbf{e}_2$$

$${}^t \mathbf{e}_2 \mathbf{e}_2 = {}^t \tilde{\mathbf{e}}_3 \tilde{\mathbf{e}}_3 + {}^t \mathbf{e}_3 \mathbf{e}_3$$

$${}^t \mathbf{e}_1 \mathbf{e}_1 = {}^t \tilde{\mathbf{e}}_2 \tilde{\mathbf{e}}_2 + {}^t \tilde{\mathbf{e}}_3 \tilde{\mathbf{e}}_3 + {}^t \mathbf{e}_3 \mathbf{e}_3$$

$${}^t \mathbf{e}_1 \mathbf{e}_1 = {}^t \tilde{\mathbf{e}}_2 \tilde{\mathbf{e}}_2 + {}^t \tilde{\mathbf{e}}_3 \tilde{\mathbf{e}}_3 + {}^t \tilde{\mathbf{e}}_4 \tilde{\mathbf{e}}_4 + \dots + {}^t \tilde{\mathbf{e}}_p \tilde{\mathbf{e}}_p + {}^t \hat{\mathbf{e}} \hat{\mathbf{e}}$$

$$\tilde{\mathbf{y}}_2 = \tilde{\mathbf{y}}_1 + \tilde{\mathbf{e}}_2 = \tilde{\mathbf{y}}_1 + \mathbf{H}_2 \mathbf{e}_1$$

$$\tilde{\mathbf{y}}_3 = \tilde{\mathbf{y}}_2 + \tilde{\mathbf{e}}_3 = \tilde{\mathbf{y}}_1 + \tilde{\mathbf{e}}_2 + \tilde{\mathbf{e}}_3$$

$${}^t \tilde{\mathbf{y}}_2 \tilde{\mathbf{y}}_2 = {}^t \tilde{\mathbf{y}}_1 \tilde{\mathbf{y}}_1 + {}^t \tilde{\mathbf{e}}_2 \tilde{\mathbf{e}}_2$$

$${}^t \tilde{\mathbf{y}}_3 \tilde{\mathbf{y}}_3 = {}^t \tilde{\mathbf{y}}_2 \tilde{\mathbf{y}}_2 + {}^t \tilde{\mathbf{e}}_3 \tilde{\mathbf{e}}_3$$

$${}^t \tilde{\mathbf{y}}_4 \tilde{\mathbf{y}}_4 = {}^t \tilde{\mathbf{y}}_3 \tilde{\mathbf{y}}_3 + {}^t \tilde{\mathbf{e}}_4 \tilde{\mathbf{e}}_4$$

$${}^t \tilde{\mathbf{y}}_4 \tilde{\mathbf{y}}_4 = {}^t \tilde{\mathbf{y}}_2 \tilde{\mathbf{y}}_2 + {}^t \tilde{\mathbf{e}}_3 \tilde{\mathbf{e}}_3 + {}^t \tilde{\mathbf{e}}_4 \tilde{\mathbf{e}}_4$$

$${}^t \tilde{\mathbf{y}}_4 \tilde{\mathbf{y}}_4 = {}^t \tilde{\mathbf{y}}_1 \tilde{\mathbf{y}}_1 + {}^t \tilde{\mathbf{e}}_2 \tilde{\mathbf{e}}_2 + {}^t \tilde{\mathbf{e}}_3 \tilde{\mathbf{e}}_3 + {}^t \tilde{\mathbf{e}}_4 \tilde{\mathbf{e}}_4$$

$${}^t \hat{\mathbf{y}} \hat{\mathbf{y}} = {}^t \tilde{\mathbf{y}}_p \tilde{\mathbf{y}}_p = {}^t \tilde{\mathbf{y}}_1 \tilde{\mathbf{y}}_1 + {}^t \tilde{\mathbf{e}}_2 \tilde{\mathbf{e}}_2 + {}^t \tilde{\mathbf{e}}_3 \tilde{\mathbf{e}}_3 + {}^t \tilde{\mathbf{e}}_4 \tilde{\mathbf{e}}_4 + \dots + {}^t \tilde{\mathbf{e}}_p \tilde{\mathbf{e}}_p$$

On a aussi

$${}^t \mathbf{y} \mathbf{y} = {}^t \tilde{\mathbf{y}}_1 \tilde{\mathbf{y}}_1 + {}^t \mathbf{e}_1 \mathbf{e}_1 = {}^t \tilde{\mathbf{y}}_2 \tilde{\mathbf{y}}_2 + {}^t \mathbf{e}_2 \mathbf{e}_2 = \dots = {}^t \tilde{\mathbf{y}}_{p-1} \tilde{\mathbf{y}}_{p-1} + {}^t \mathbf{e}_{p-1} \mathbf{e}_{p-1} = {}^t \hat{\mathbf{y}} \hat{\mathbf{y}} + {}^t \hat{\mathbf{e}} \hat{\mathbf{e}} = n - 1$$

ANNEXE 6

Connaissant \mathbf{X} , \mathbf{y} et les \mathbf{t}_i on peut calculer toutes les réponses solutions.

Première réponse

$$\mathbf{H}_1 = \mathbf{X} \mathbf{t}_1 \left[{}^t \mathbf{t}_1 {}^t \mathbf{X} \mathbf{X} \mathbf{t}_1 \right]^{-1} {}^t \mathbf{t}_1 {}^t \mathbf{X}$$

$$\tilde{\mathbf{y}}_1 = \mathbf{H}_1 \mathbf{y}$$

$$\mathbf{e}_1 = (\mathbf{I} - \mathbf{H}_1) \mathbf{y}$$

Deuxième réponse

$$\mathbf{X}_{r_1} = (\mathbf{I} - \mathbf{H}_1) \mathbf{X}$$

$$\mathbf{X}_{r_1} = \left[\mathbf{I} - \mathbf{X} \mathbf{t}_1 \left({}^t \mathbf{t}_1 {}^t \mathbf{X} \mathbf{X} \mathbf{t}_1 \right)^{-1} {}^t \mathbf{t}_1 {}^t \mathbf{X} \right] \mathbf{X}$$

$$\tilde{\mathbf{e}}_2 = \mathbf{H}_2 \mathbf{e}_1$$

$$\tilde{\mathbf{y}}_2 = \tilde{\mathbf{y}}_1 + \tilde{\mathbf{e}}_2 = \tilde{\mathbf{y}}_1 + \mathbf{H}_2 \mathbf{e}_1$$

$$\mathbf{e}_2 = (\mathbf{I} - \mathbf{H}_2) \mathbf{e}_1$$

\mathbf{H}_2 peut être exprimé avec \mathbf{X} , \mathbf{y} et les \mathbf{t}_i

Troisième réponse

$$\mathbf{X}_{r_2} = (\mathbf{I} - \mathbf{H}_2) (\mathbf{I} - \mathbf{H}_1) \mathbf{X}$$

$$\mathbf{H}_3 = \mathbf{X}_{r_2} \mathbf{t}_3 \left[{}^t \mathbf{t}_3 {}^t \mathbf{X}_{r_2} \mathbf{X}_{r_2} \mathbf{t}_3 \right]^{-1} {}^t \mathbf{t}_3 {}^t \mathbf{X}_{r_2}$$

$$\tilde{\mathbf{e}}_3 = \mathbf{H}_3 \mathbf{e}_2$$

$$\tilde{\mathbf{y}}_3 = \tilde{\mathbf{y}}_2 + \tilde{\mathbf{e}}_3 = \tilde{\mathbf{y}}_1 + \tilde{\mathbf{e}}_2 + \tilde{\mathbf{e}}_3$$

$$\mathbf{e}_3 = (\mathbf{I} - \mathbf{H}_3) \mathbf{e}_2$$

\mathbf{H}_3 peut être exprimé avec \mathbf{X} , \mathbf{y} et les \mathbf{t}_i

k ième réponse

$$\mathbf{X}_{r_{k-1}} = (\mathbf{I} - \mathbf{H}_{k-1}) \cdots (\mathbf{I} - \mathbf{H}_2) (\mathbf{I} - \mathbf{H}_1) \mathbf{X}$$

$$\mathbf{H}_k = \mathbf{X}_{r_{k-1}} \mathbf{t}_k \left[{}^t \mathbf{t}_k {}^t \mathbf{X}_{r_{k-1}} \mathbf{X}_{r_{k-1}} \mathbf{t}_k \right]^{-1} {}^t \mathbf{t}_k {}^t \mathbf{X}_{r_{k-1}}$$

$$\tilde{\mathbf{e}}_k = \mathbf{H}_k \mathbf{e}_{k-1}$$

$$\tilde{\mathbf{y}}_k = \tilde{\mathbf{y}}_{k-1} + \tilde{\mathbf{e}}_k = \tilde{\mathbf{y}}_1 + \tilde{\mathbf{e}}_2 + \tilde{\mathbf{e}}_3 + \cdots + \tilde{\mathbf{e}}_k$$

$$\mathbf{e}_k = (\mathbf{I} - \mathbf{H}_k) \mathbf{e}_{k-1}$$

\mathbf{H}_k peut être exprimé avec \mathbf{X} , \mathbf{y} et les \mathbf{t}_i

ANNEXE 7

Les grandeurs \mathbf{X}_{r_i} et \mathbf{e}_k suivent la même évolution au cours des calculs. En effet, on a

$$\mathbf{X} \mathbf{t}_i^{(j-1)'} = (\mathbf{I} - \mathbf{H}_{j-1}) \cdots (\mathbf{I} - \mathbf{H}_2) (\mathbf{I} - \mathbf{H}_1) \mathbf{X} \mathbf{t}_i$$

$$\mathbf{X}_{r_{j-1}} = (\mathbf{I} - \mathbf{H}_{j-1}) \cdots (\mathbf{I} - \mathbf{H}_2) (\mathbf{I} - \mathbf{H}_1) \mathbf{X}$$

$$\mathbf{e}_k = (\mathbf{I} - \mathbf{H}_k) \cdots (\mathbf{I} - \mathbf{H}_2) (\mathbf{I} - \mathbf{H}_1) \mathbf{e}_1$$

Avec les matrices de projection, on a

$$\tilde{\mathbf{e}}_i = \mathbf{H}_i \cdots \mathbf{H}_2 \mathbf{H}_1 \mathbf{e}_1$$

ANNEXE 8

Si l'on note \mathbf{t}_p les vecteurs \mathbf{t} mesurés dans l'espace des réponses possibles ERP1 de p dimensions, la norme de ces vecteurs est égale à

$$\|\mathbf{t}_p\| = \sqrt{{}^t\mathbf{t}_p \mathbf{t}_p}$$

Si l'on note \mathbf{t}_n les vecteurs \mathbf{t} mesurés dans le référentiel des réponses de n dimensions, on a

$$\mathbf{t}_n = \mathbf{X}\mathbf{t}_n$$

Par conséquent, la norme de ces vecteurs est égale à

$$\|\mathbf{t}_n\| = \sqrt{{}^t\mathbf{t}_n {}^t\mathbf{X} \mathbf{X}\mathbf{t}_n}$$

On a donc

$${}^t\mathbf{t}_p \mathbf{t}_p = {}^t\mathbf{t}_n {}^t\mathbf{X} \mathbf{X}\mathbf{t}_n$$

La matrice ${}^t\mathbf{X} \mathbf{X}$ est, à un facteur $n-1$ près, égale au tenseur fondamental ou tenseur métrique de l'espace des réponses possibles.

$$\mathbf{G} = \begin{bmatrix} \frac{{}^t\mathbf{x}_1 \mathbf{x}_1}{\|\mathbf{x}_1\| \|\mathbf{x}_1\|} & \frac{{}^t\mathbf{x}_1 \mathbf{x}_2}{\|\mathbf{x}_1\| \|\mathbf{x}_2\|} \\ \frac{{}^t\mathbf{x}_1 \mathbf{x}_2}{\|\mathbf{x}_1\| \|\mathbf{x}_2\|} & \frac{{}^t\mathbf{x}_2 \mathbf{x}_2}{\|\mathbf{x}_2\| \|\mathbf{x}_2\|} \end{bmatrix}$$

En effet, les éléments du tenseur métrique sont les produits scalaires des vecteurs de base normés de la nouvelle base. Ce tenseur peut également être écrit :

$$\mathbf{G} = \begin{bmatrix} 1 & \cos(\mathbf{x}_1, \mathbf{x}_2) \\ \cos(\mathbf{x}_1, \mathbf{x}_2) & 1 \end{bmatrix}$$

On reconnaît ici la matrice de variance covariance des vecteurs normés.

ANNEXE 9

La régression PLS1 est illustrée par le fichier Excel « PLS1_4_coef ».

Le lecteur pourra entrer les 4 coefficients d'un modèle et retrouver les 3 solutions PLS ainsi que la solution classique de la régression. Il pourra constater que la quatrième solution PLS1 est égale à la solution classique.

ANNEXE 10

La RLSO (Régression Linéaire Séquentielle Orthogonale) est illustrée par le fichier Excel « RLSO_4_coef ».

Le lecteur pourra entrer un modèle de 4 coefficients et modifier à sa guise les directions de projections. Le logiciel calculera les 3 solutions RLSO ainsi que la solution classique de la régression. Il pourra constater que la quatrième solution RLSO est toujours égale à la solution classique.

Lorsque les directions de projection sont celles de la régression PLS1, la RLSO fournit les mêmes solutions que la régression PLS1.

Ces exemples sous Excel ont bénéficié de l'aide de Henry Aubert et de Jacques Vaillé.

Bibliographie

Régression classique

1. DRAPER Norman and SCHMITT H
"Applied Regression Analysis"
John Wiley and Sons. New-York. 708 pages. (1981).
2. BATES Douglas M. and WATTS Donald G.
"Nonlinear Regression analysis and its applications"
John Wiley and Sons. New-York. 365 pages. (1988).
3. WILLIAMS E.J.
"Regression Analysis"
John Wiley and Sons. New-York. (1967).

Régression PLS

4. TENENHAUS Michel
"La régression PLS. Théorie et pratique"
Editions Technip. 254 pages (1998.)Paris. ISBN 2-7108-0735
5. WOLD H.
"Soft modeling. The basic design and some extensions"
In vol II of Joreskog K.G.and Wold H Eds.
System under indirect observation. North Holland, Amsterdam. (1982).
6. WOLD Swante, RUHE A., DUNN III W. J. and WOLD H.
"The collinearity problem in Linear Regression. The Partial Least Square Approach to Generalized Inverses"
SIAM J. Sci. Stat. Comp. 5 (1984). p. 735-743.
7. HOSKULDSSONN A.
"PLS Regression Methods"
J. Chemometrics 2, p. 211-228 (1988).
8. TENENHAUS M., GAUCHI Jean-Pierre et MENARDO C.
"Régression PLS et applications"
Rev. Statistique Appliquée vol. **XLIII** (1), p.7-63. (1995).
9. GAUCHI Jean-Pierre
"Utilisation de la régression PLS pour l'analyse des plans d'expériences en chimie de formulation"
Rev. Statistique Appliquée vol. **XLIII** (1), p.65-89. (1995).
10. BRANDVIK P.J. and DALING P.S.
"Optimising oil spill dispersants as a function of oil type and weathering degree: a multivariate approach using partial least squares (PLS)."
Chemometrics and Intelligent Laboratory Systems 42 (1998): 73-91.
11. LINDGREN F., GELADI P. WOLD S.
"The kernel algorithm for PLS"
J. Chemometrics. 7, p. 45-59. (1993).
12. CHAVENT Marie et PATOUILLE Brigitte.
"Calcul des coefficients de régression et du Press en régression PLS1"
Modulad n° 30. p. 1-11. (2003).
13. BASTIEN Philippe, VINZI Vincenzo Esposito, TENENHAUS Michel
"PLS generalised linear regression"
Computational statistics and data analysis. 48, p. 17-46. (2005).