



**HAL**  
open science

## Plans d'expériences: Plans associés à la régression linéaire

Charlie Sire

► **To cite this version:**

Charlie Sire. Plans d'expériences: Plans associés à la régression linéaire. École d'ingénieur. France. 2023. hal-04527311

**HAL Id: hal-04527311**

**<https://hal.science/hal-04527311v1>**

Submitted on 2 Apr 2024

**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.

# Plans d'expériences

## Partie 1 : Plans associés à la regression linéaire

Charlie SIRE

6 décembre 2023

- 1 Rappels sur la régression linéaire
- 2 Introduction aux plans d'expériences
- 3 Les plans d'expériences de criblage
- 4 Plans de surface de réponse
  - Les plans classiques
  - Les plans optimisés

Contexte et notations :

- $x_1, \dots, x_p$  sont des variables explicatives (descripteurs)
- $Y$  est la variable (quantitative) à expliquer (variable d'intérêt).  
Objectif : on cherche à approcher la relation entre  $Y$  et  $x_1, \dots, x_p$  à partir d'un échantillon à  $n$  observations.

# Hypothèses

On suppose que la sortie est la réalisation du modèle suivant :

$$Y_i = \beta_0 + \sum_{j=1}^{j=p} \beta_j x_{ij} + \mathcal{E}_i, \quad \forall 1 \leq i \leq n$$

# Hypothèses

On suppose que la sortie est la réalisation du modèle suivant :

$$Y_i = \beta_0 + \sum_{j=1}^{j=p} \beta_j x_{ij} + \mathcal{E}_i, \quad \forall 1 \leq i \leq n$$

avec :

- $\beta_0, \dots, \beta_p$  nombres réels déterministes inconnus
- $\mathcal{E}_1, \dots, \mathcal{E}_n$   $n$  variables aléatoires :
  - non corrélées
  - de moyenne nulle  $\mathbb{E}[\mathcal{E}_i] = 0$
  - de variance constante  $\mathbb{V}(\mathcal{E}_i) = \sigma^2$

Rq :  $\mathbf{x}_j$  peut être de la forme  $\mathbf{x}_1 \mathbf{x}_2$  ou  $\mathbf{x}_1^2$  par exemple

# Forme matricielle

Le modèle précédent s'écrit matriciellement :

$$\begin{bmatrix} Y_1 \\ \cdot \\ \cdot \\ Y_n \end{bmatrix} = \mathbf{x} \begin{bmatrix} \beta_0 \\ \cdot \\ \cdot \\ \beta_p \end{bmatrix} + \begin{bmatrix} \mathcal{E}_1 \\ \cdot \\ \cdot \\ \mathcal{E}_n \end{bmatrix}$$

$$\text{avec } \mathbf{x} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix},$$

$$\mathbb{E}\left(\begin{bmatrix} \mathcal{E}_1 \\ \cdot \\ \cdot \\ \mathcal{E}_n \end{bmatrix}\right) = \mathbf{0}_n, \text{ et } \mathbb{V}\left(\begin{bmatrix} \mathcal{E}_1 \\ \cdot \\ \cdot \\ \mathcal{E}_n \end{bmatrix}\right) = \sigma^2 Id_n$$

# Estimateur des moindres carrés

Le vecteur  $\hat{\beta}$  qui minimise la fonction

$$C(\beta) = \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}))^2$$

est  $\hat{\beta} = ({}^t\mathbf{xx})^{-1}({}^t\mathbf{xy})$

On va donc considérer l'estimateur des moindres carrés :

$$\hat{\mathbf{B}} = ({}^t\mathbf{xx})^{-1}({}^t\mathbf{xY})$$

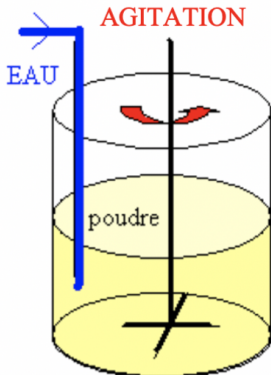
On a de plus :  $\mathbb{V}(\hat{\mathbf{B}}) = \sigma^2({}^t\mathbf{xx})^{-1}$



- 1 Rappels sur la régression linéaire
- 2 Introduction aux plans d'expériences**
- 3 Les plans d'expériences de criblage
- 4 Plans de surface de réponse
  - Les plans classiques
  - Les plans optimisés

# Exemple introductif

- 1 On humidifie et on agite une poudre au sein d'un granulateur.
- 2 Des agglomérats appelés granules se forment alors, sous l'effet de l'agitation, de l'humidité et de la température.
- 3 Un temps de repos permet de finir le grossissement des granules.
- 4 On s'intéresse au final à la taille moyenne des granules présents dans le granulateur



## Variables explicatives :

- taux de remplissage du granulateur (quantité de poudre introduite initialement, en %)
- débit d'eau à introduire dans le granulateur (en  $mm^3s^{-1}$ )
- le temps de mouillage (en  $s$ )
- la température de maintien (en  $^{\circ}C$ )
- la vitesse d'agitation de la pale (en  $toursmn^{-1}$ )
- le temps d'agitation (en  $s$ )
- le temps de repos (en  $s$ ).

**Variable à expliquer :** Taille moyenne des granules

# Espace des variables explicatives

<b>Facteurs</b>	<b>min</b>	<b>moyenne</b>	<b>max</b>	<b>unité</b>
<b>débit d'eau</b>	30	60	90	mm <sup>3</sup> s <sup>-1</sup>
<b>température de maintien</b>	20	25	30	°C
<b>taux de remplissage</b>	0,74	0,78	0,82	%
<b>temps de mouillage</b>	200	500	800	s
<b>temps de repos</b>	0	200	400	s
<b>vitesse agitation</b>	148	248	348	tours mn-1
<b>temps d'agitation</b>	3000	4000	5000	s

Les plans d'expériences interviennent lorsque le nombre d'essai (budget, temps etc) est limité et les facteurs sont contrôlés.

Il peuvent permettre de :

- Identifier les facteurs les plus influents (Plans de criblage)
- Prédire la taille des granules sur l'ensemble du domaine (Plans de surface de réponses)

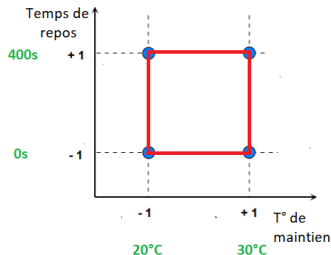
- 1 Rappels sur la régression linéaire
- 2 Introduction aux plans d'expériences
- 3 Les plans d'expériences de criblage
- 4 Plans de surface de réponse
  - Les plans classiques
  - Les plans optimisés

On se place dans le cas d'un modèle de régression de degré 1 avec ou sans interaction :

$$Y = \beta_0 + \sum \beta_i x_i + \sum \beta_{ij} x_i x_j + \dots + \beta_{ij\dots p} x_i x_j \dots x_p + \varepsilon$$

Chaque variable est normalisée entre -1 et 1 :

$$\tilde{x} = 2 \frac{x - x_{min}}{x_{max} - x_{min}} - 1$$



**Question:** Combien de  $y$  a-t-il de coefficients  $\beta_0, \dots, \beta_{1,2,\dots,p}$  ?

# Plan factoriel à 2 niveau

## Principe

Les plans d'expérience factoriels permettent de traiter un nombre  $p$  quelconque de facteurs à 2 niveaux (facteur quantitatif) ou 2 modalités (facteur qualitatif).

Les expériences à réaliser le sont pour les niveaux -1 et 1 : niveaux "haut" et "bas"

## Orthogonalité

Pour évaluer l'effet d'un facteur  $x_j$ , il faut que pour chaque niveau de ce facteur, tous les autres facteurs apparaissent autant de fois au niveau "haut" qu'au niveau "bas", sinon leur effet biaiserait celui de  $x_j$ .  $n$  doit donc être une puissance de 2.

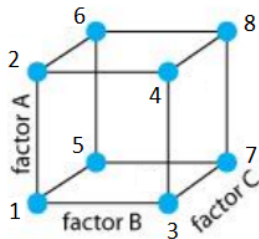
Le plan d'expérience est alors dit orthogonal :  $({}^t\mathbf{xx})^{-1}$  est proportionnel à l'identité



# Plan factoriel complet $2^p$

Il consiste à effectuer les  $n = 2^p$  expériences possibles. Par exemple, pour  $p=3$  avec 3 variables a, b et c :

EXP	A	B	C
1	-1	-1	-1
2	+1	-1	-1
3	-1	+1	-1
4	+1	+1	-1
5	-1	-1	+1
6	+1	-1	+1
7	-1	+1	+1
8	+1	+1	+1



# Ajout de paramètres estimables

Le plan factoriel complet nous donne un système  $n = 2^p$  équations.

On peut donc inclure les paramètres d'interaction, et estimer  $2^p$  paramètres :  $\beta_0, \beta_i, \beta_{ij}, \beta_{ij..p}$

Dans le cas  $p = 3$ , on obtient :

EXP	A	B	C	A*B	B*C	A*C	A*B*C
1	-1	-1	-1	+1	+1	+1	-1
2	+1	-1	-1	-1	+1	-1	+1
3	-1	+1	-1	-1	-1	+1	+1
4	+1	+1	-1	+1	-1	-1	-1
5	-1	-1	+1	+1	-1	-1	+1
6	+1	-1	+1	-1	-1	+1	-1
7	-1	+1	+1	-1	+1	-1	-1
8	+1	+1	+1	+1	+1	+1	+1

# Matrice X

La matrice  $\mathbf{x}$  s'obtient en rajoutant une colonne de 1 :

$$\mathbf{x} = \begin{bmatrix} +1 & -1 & -1 & -1 & +1 & +1 & +1 & -1 \\ +1 & +1 & -1 & -1 & -1 & +1 & -1 & +1 \\ +1 & -1 & +1 & -1 & -1 & -1 & +1 & +1 \\ +1 & +1 & +1 & -1 & +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 \\ +1 & +1 & -1 & +1 & -1 & -1 & +1 & -1 \\ +1 & -1 & +1 & +1 & -1 & +1 & -1 & -1 \\ +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \end{bmatrix}$$

Et on a :  $\hat{\mathbf{B}} = ({}^t\mathbf{xx})^{-1}({}^t\mathbf{xY})$

# Exercice

- 1 Vérifier que les vecteurs de  $\mathbf{x}$  sont orthogonaux
- 2 Quelle est la conséquence sur  ${}^t\mathbf{xx}$  ?
- 3 En déduire les  $\hat{\beta}_0, \hat{\beta}_A, \hat{\beta}_B, \hat{\beta}_C, \hat{\beta}_{AB}, \hat{\beta}_{AC}, \hat{\beta}_{BC}, \hat{\beta}_{ABC}$  pour

$$\mathbf{y} = \begin{bmatrix} 5 \\ 10 \\ 5 \\ 20 \\ 2 \\ 3 \\ 20 \\ 15 \end{bmatrix}$$

- 4 Classer les facteurs par ordre d'influence

On a  ${}^t\mathbf{xx} = 2^p I_d_n$

$$\text{Donc } \hat{\beta}_i = \frac{1}{2^p} (\sum_{u|x_{u,i}=+1} y_u - \sum_{v|x_{v,i}=-1} y_v)$$

Formulé autrement, l'effet d'un facteur (ou interaction) est égal (à un facteur près) aux différences des moyennes de la réponse entre le niveau 1 et -1 de ce facteur (ou interaction)

# Plans factoriels fractionnaires $2^{p-q}$

Les plans factoriels fractionnaires sont des plans factoriels qui permettent d'étudier tous les facteurs mais dont le nombre d'essais est réduit par rapport aux plans factoriels complets. Un plan factoriel fractionnaire à 2 fois moins, ou 4 fois moins ou  $2^q$  fois moins d'essais que le factoriel complet correspondant.

On appellera plan  $2^{p-q}$  un plan comprenant  $2^{p-q}$  expériences pour  $p$  facteurs à deux niveaux.

On a alors  $n = 2^{p-q}$  équations et  $2^p$  inconnues si l'on veut prendre en compte l'ensemble des interactions.

# Illustration confusion

Dans l'exemple précédent avec  $p = 3$ , si on ne pouvait réaliser que les 4 premières expériences :

<b>EXP</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>A*B</b>	<b>B*C</b>	<b>A*C</b>	<b>A*B*C</b>
1	-1	-1	-1	+1	+1	+1	-1
2	+1	-1	-1	-1	+1	-1	+1
3	-1	+1	-1	-1	-1	+1	+1
4	+1	+1	-1	+1	-1	-1	-1

# Illustration confusion

Dans l'exemple précédent avec  $p = 3$ , si on ne pouvait réaliser que les 4 premières expériences :

EXP	A	B	C	A*B	B*C	A*C	A*B*C
1	-1	-1	-1	+1	+1	+1	-1
2	+1	-1	-1	-1	+1	-1	+1
3	-1	+1	-1	-1	-1	+1	+1
4	+1	+1	-1	+1	-1	-1	-1



## Construction d'un plan $2^{p-q}$

On peut construire un tel plan en construisant d'abord un plan factoriel complet pour  $(p-q)$  facteurs puis on définit les  $q$  autres facteurs en fonction des  $(p-q)$  par produit uniquement.

Il va y avoir alors confusion entre certains facteurs et certaines interactions.

On appelle clé du plan l'ensemble des relations que l'on exprime sous la forme  $1 = \dots$

En effet,  $F = A.B$ , s'écrit  $F.A.B = 1$  car  
 $F = A.B \Rightarrow F.F = 1 = F.A.B$

La clé indique à elle-seule l'ensemble des confusions.

# Exemple : plan $2^{5-2}$ pour 5 facteurs A,B,C,D et E

Plan complet pour A, B et C puis on ajoute D et E avec  $D=AB$  et  $E = AC$

EXP	A	B	C	D = A*B	E = A*C
1	-1	-1	-1	+1	+1
2	+1	-1	-1	-1	-1
3	-1	+1	-1	-1	+1
4	+1	+1	-1	+1	-1
5	-1	-1	+1	+1	-1
6	+1	-1	+1	-1	+1
7	-1	+1	+1	-1	-1
8	+1	+1	+1	+1	+1

Quelle est la clé du plan ?

# Exemple : plan $2^{5-2}$ pour 5 facteurs A,B,C,D et E

Plan complet pour A, B et C puis on ajoute D et E avec  $D=AB$  et  $E = AC$

EXP	A	B	C	D = A*B	E = A*C
1	-1	-1	-1	+1	+1
2	+1	-1	-1	-1	-1
3	-1	+1	-1	-1	+1
4	+1	+1	-1	+1	-1
5	-1	-1	+1	+1	-1
6	+1	-1	+1	-1	+1
7	-1	+1	+1	-1	-1
8	+1	+1	+1	+1	+1

Quelle est la clé du plan ?  $1 = ADB = ACE$

# Calcul des effets des facteurs

L'équation de régression linéaire à  $2^5$  inconnues (avec les interactions) se réécrit avec  $2^3$  inconnues :

$$y = \alpha_0 + \alpha_A \times A + \alpha_B \times B + \alpha_C \times C + \alpha_D \times D + \alpha_E \times E + \alpha_{BC} \times BC + \alpha_{CD} \times CD + \epsilon$$

avec :

$$\alpha_0 = \beta_0 + \beta_{ABD} + \beta_{ACE} + \beta_{BCDE}$$

$$\alpha_A = \beta_A + \beta_{BD} + \beta_{CE} + \beta_{ABCDE}$$

$$\alpha_B = \beta_B + \beta_{AD} + \beta_{CDE} + \beta_{ABCE}$$

$$\alpha_C = \beta_C + \beta_{AE} + \beta_{BDE} + \beta_{ABCD}$$

$$\alpha_D = \beta_D + \beta_{AB} + \beta_{BCE} + \beta_{ACDE}$$

$$\alpha_E = \beta_E + \beta_{AC} + \beta_{BCD} + \beta_{ABDE}$$

$$\alpha_{BC} = \beta_{BC} + \beta_{DE} + \beta_{ABE} + \beta_{ACD}$$

$$\alpha_{CD} = \beta_{CD} + \beta_{BE} + \beta_{ABC} + \beta_{ADE}$$

Les  $\alpha_i$  sont appelés les contrastes du modèle.

Alias de A :  $A = BD = CE = ABCDE$

Les alias peuvent se déduire de la clé (et inversement)

# Hypothèses d'interprétation

Le fait que certains effets soient confondus avec d'autres complique l'interprétation des contrastes calculés. Des hypothèses vont alors être formulées :

- H1 : Interactions entre 3 facteurs ou plus sont négligeables
- H2 : Contraste nul  $\Rightarrow$  effets reliés tous nuls

Rappel : L'objectif des plans factoriels est d'étudier l'influence des variables explicatives

# Résolution

On appelle résolution du plan noté  $\rho$  l'entier :

$$\rho = \inf\{\text{nb de symboles des éléments de l'alias de 1}\}$$

Dans l'exemple précédent, le plan est de résolution III :

$$1 = ADB = ACE = BDCE$$

On a ainsi, par exemple :

- Résolution III : certains effets principaux sont confondus avec des interactions "double" (produit de 2 facteurs)
- Résolution IV : un effet principal ne peut être confondu avec une interaction "double", mais certaines interactions "doubles" sont confondues entre elles
- Résolution V : on peut poser un modèle avec toutes les interactions "doubles" et les effets principaux sans confusion

## Remarque : Estimation de l'écart type

Dans tout ce qui précède on ne voulait estimer que les effets des facteurs  $\hat{\beta}_i$ .

Si l'on veut estimer l'écart type  $\sigma$ , cela rajoute une inconnue :

$$\hat{\sigma} = \frac{1}{\sqrt{n - \tilde{p} - 1}} \|\mathbf{y} - \hat{\mathbf{y}}\|$$

avec  $n = \text{nb d'équations (nb de lignes de X)}$

$\tilde{p} + 1 = \text{nb de facteurs pris en compte (nb de colonnes de X)}$

⇒ Si on veut estimer la variance,  $\mathbf{x}$  ne doit pas être une matrice carrée

Par exemple dans le plan complet  $2^p$ , au lieu d'estimer l'ensemble des interactions on peut négliger celle de plus haut "degré"  $\beta_{12..p}$  pour gagner un degré de liberté :  $\tilde{p} = 2^p - 1 = n - 1 \Rightarrow$  on peut estimer  $\sigma$

## Conclusion partie 3

Plans factoriels à 2 niveaux sont :

- Simples à construire
- Peu coûteux en essais
- Orthogonaux

Mais : Ils ne concernent que les modèles de degré 1, et ne permettent donc pas l'étudier la surface de réponse lorsque l'on suspecte une courbure



# Content

- 1 Rappels sur la régression linéaire
- 2 Introduction aux plans d'expériences
- 3 Les plans d'expériences de criblage
- 4 Plans de surface de réponse
  - Les plans classiques
  - Les plans optimisés

# Exercice

Exercice On veut construire un plan pour 1 facteur quantitatif  $A$  appartenant à  $[-1, 1]$ . On souhaite estimer un modèle de degré 2.

- Combien de paramètres y-a-t-il à estimer? Cela conditionne le nombre d'essais à effectuer au minimum.
- Est-ce possible d'estimer un modèle de degré 2 avec des expériences dupliquées en  $-1$  et en  $1$ ?
- On choisit de réaliser 3 expériences : 1 en  $-1$ , 1 en  $1$ , 1 en  $a$ ,  $a \in [-1, 1]$ . A votre avis, comment choisir  $a$  pour une estimation la plus précise possible des paramètres?

# Content

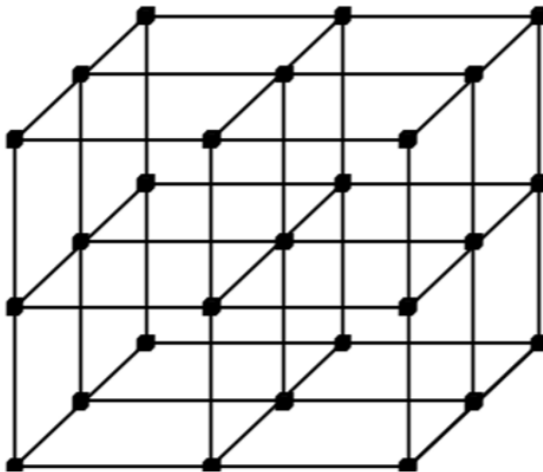
- 1 Rappels sur la régression linéaire
- 2 Introduction aux plans d'expériences
- 3 Les plans d'expériences de criblage
- 4 Plans de surface de réponse
  - Les plans classiques
  - Les plans optimisés

# Les plans classiques

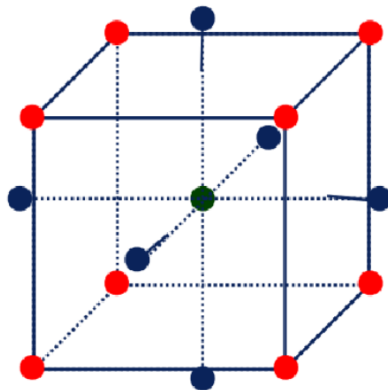
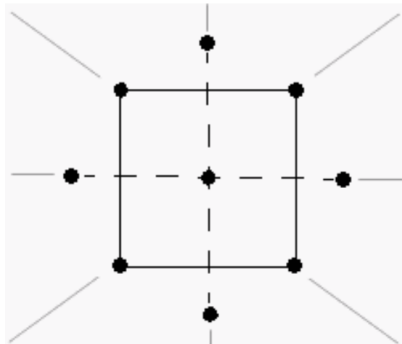
Présentations des quelques plans de surface de réponse classiques :

- ① Plan factoriel complet à 3 niveaux
- ② Plan composite centré
- ③ Plan de Box Benkhen

# Le plan factoriel complet à 3 niveaux



# Plan composite centré

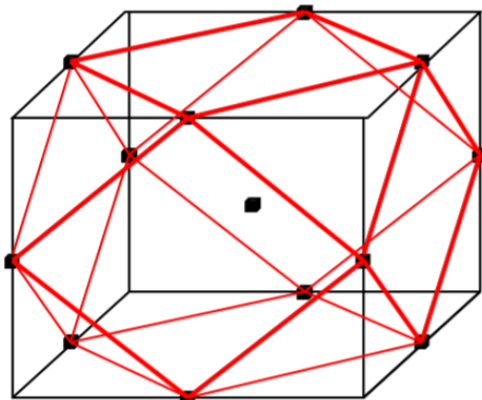


Principe : plan factoriel + points axiaux et centraux

Avantage : Prend facilement la suite d'un plan factoriel

Inconvénient : Point à l'extérieur du cube

# Le plan de Box Benkhen



Principe : Points centraux des bords

Avantage : Moins de points que les plans composites

Inconvénient : Pas de plan factoriel imbriqué donc inapproprié pour les expériences séquentielles

# Content

- 1 Rappels sur la régression linéaire
- 2 Introduction aux plans d'expériences
- 3 Les plans d'expériences de criblage
- 4 Plans de surface de réponse
  - Les plans classiques
  - Les plans optimisés



# Les plans optimisés

On a :

$$\mathbb{V}(\hat{\mathbf{B}}) = \sigma^2({}^t\mathbf{xx})^{-1}$$

Donc

$$\forall 1 \leq i \leq n, \mathbb{V}(\hat{B}_i) = \sigma^2({}^t\mathbf{xx})_{ii}^{-1}$$

⇒ On s'intéresse à la matrice  $({}^t\mathbf{xx})^{-1}$  (et donc à  $({}^t\mathbf{xx})$ ) pour minimiser la variance.

Dans la suite, on note  $\lambda_0, \dots, \lambda_p$  les valeurs propres de  $({}^t\mathbf{xx})^{-1}$ .

## D-optimalité :

$$({}^t\mathbf{xx})^{-1} = \frac{1}{\det({}^t\mathbf{xx})} \times {}^t [({}^t\mathbf{xx})^*]$$

où  $({}^t\mathbf{xx})^*$  désigne la matrice des cofacteurs associée à la matrice  ${}^t\mathbf{xx}$ .

Donc grossièrement, les éléments de  $\mathbb{V}(\hat{\mathbf{B}})$  (et les diagonaux notamment) seront donc d'autant plus petits que  $\det({}^t\mathbf{xx})$  est grand.

On cherche alors la position des points  $\pi_D$  :

$$\pi_D = \operatorname{argmax}(\det({}^t\mathbf{xx})) = \operatorname{argmin}(\lambda_0 \times \dots \times \lambda_p)$$

# A-optimalité

On a  $\sum_i \mathbb{V}(\hat{B}_i) = \sigma^2 \text{Trace}(({}^t\mathbf{X}\mathbf{X})^{-1}) = \sigma^2 \sum_i \lambda_i$

On cherche alors la position des points  $\pi_A$  :

$$\pi_A = \underset{i}{\operatorname{argmin}}(\text{Trace}(({}^t\mathbf{X}\mathbf{X})^{-1})) = \underset{i}{\operatorname{argmin}}(\sum_i \lambda_i)$$

On peut également s'intéresser directement à la variance de prédiction, en cherchant la position des points  $\pi_G$  telle que :

$$\pi_G = \operatorname{argmin} [ \max_{\mathbf{x}_{new} \in \mathcal{X}} (\mathbb{V}(\hat{Y}(\mathbf{x}_{new}))) ]$$

avec  $\mathbb{V}(\hat{Y}(\mathbf{x}_{new})) = \sigma^2 ({}^t \mathbf{x}_{new} ({}^t \mathbf{X} \mathbf{X})^{-1} \mathbf{x}_{new})$

# Retour sur l'exercice

On veut construire un plan pour 1 facteur quantitatif A appartenant à  $[-1, 1]$ . On souhaite estimer un modèle de degré 2. On choisit de réaliser 3 expériences : 1 en -1, 1 en 1, 1 en a,  $a \in [-1, 1]$ . Comment choisir a pour une estimation la plus précise possible des paramètres?

