

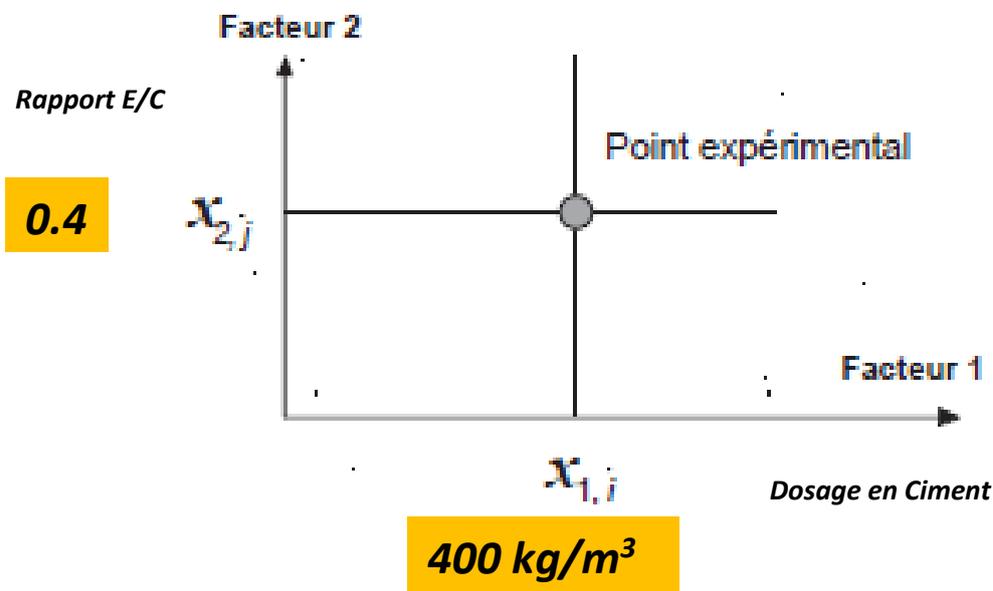
## 1.5 Points expérimentaux

Dans un espace à deux dimensions, le niveau  $i$  du facteur 1, noté  $x_{1,i}$ , et le niveau  $j$  du facteur 2, noté  $x_{2,j}$ , peuvent être considérés comme les coordonnées d'un point de l'espace expérimental ou du domaine d'étude (figure 1.6). Par exemple, si le niveau du dosage en ciment est  $400 \text{ kg/m}^3$  et celui du rapport Eau/Ciment est  $0.4$  les coordonnées du point expérimental sont :

$$x_{1,i} = 400 \text{ kg/m}^3$$

$$x_{2,j} = 0.4$$

Une expérience donnée est alors représentée par un point dans ce système d'axes. C'est la raison pour laquelle une expérience est souvent désignée par l'expression point expérimental, point d'expérience ou même simplement point. Un plan d'expériences est donc représenté par un ensemble de points expérimentaux, eux-mêmes situés dans l'espace expérimental. Dans l'exemple que nous avons pris, l'expérience est conduite sur une formulation avec un dosage en ciment est  $400 \text{ kg/m}^3$  et un rapport Eau/Ciment est  $0.4$



Jusqu'à trois facteurs, il est possible de dessiner le domaine d'étude. Au-delà de trois facteurs, on utilise une représentation en tableau, dite *matricielle*, plus générale puisqu'elle permet de représenter les points d'expériences dans un hypervolume à un nombre quelconque de dimensions.

## 1.6 Plans d'expériences

### 1.6.1 Méthodologie des plans sans contrainte

Le choix du nombre et de l'emplacement des points d'expériences est le problème fondamental des plans d'expériences. On cherche à réaliser le minimum d'expériences tout en réduisant l'influence de l'erreur expérimentale sur les modélisations mathématiques qui serviront à prendre des décisions. On atteint ce but en considérant

Les propriétés mathématiques et statistiques qui relient la réponse aux facteurs. Lorsqu'il n'y a pas de contraintes sur le domaine d'étude, il existe des plans classiques qui possèdent d'excellentes qualités statistiques et qui permettent de modéliser les réponses dans les meilleures conditions (figure 1.7). Lorsqu'il existe des contraintes, il faut construire des plans sur mesure en recherchant la position des points expérimentaux qui conduisent, là aussi, à de bonnes qualités statistiques et à une bonne modélisation des réponses.

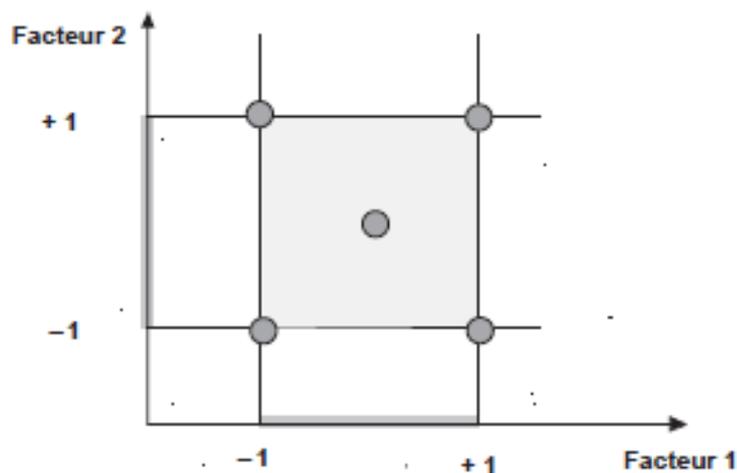


Figure 1.7 – Exemple de disposition des points expérimentaux dans un domaine sans contraintes.

### 1.6.2 Méthodologie des plans avec contrainte

La procédure de construction des plans dont le domaine est contraint est la suivante :

1. On définit le domaine de chacun des facteurs (niveau bas et niveau haut).
2. On définit les contraintes qui pèsent sur les facteurs. Ces contraintes sont exprimées par des relations d'inégalité entre les facteurs et elles définissent les zones autorisées, c'est-à-dire celles où les expériences sont possibles, et les zones interdites, c'est-à-dire celles où les expériences ne doivent pas être exécutées.

3. On définit les niveaux des facteurs les plus intéressants pour l'étude, autres que les niveaux bas et hauts. Le plus souvent entre 2 et 5 niveaux supplémentaires par rapport facteur.
4. On construit une grille en prenant en compte toutes les combinaisons des niveaux des facteurs. Cette grille ne doit contenir que les points expérimentaux réalistes, c'est-à-dire les points du domaine autorisé. Ces points constituent les points candidats (figure 1.8).

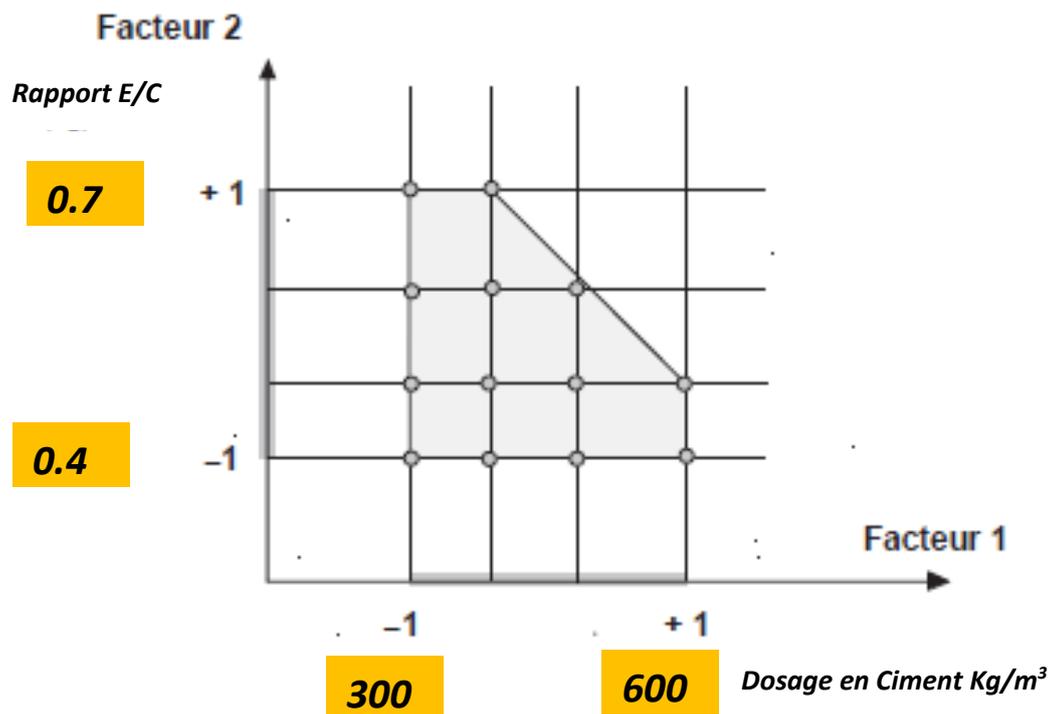


Figure 1.8 – La grille des points candidats est constituée d'expériences possibles dans le domaine d'étude.

5. On choisit une fonction reliant *a priori* la réponse aux facteurs.
6. On sélectionne, en fonction d'un critère d'optimalité choisi, le nombre et l'emplacement des points expérimentaux les plus utiles à la modélisation du phénomène étudié (figure 1.9). Cette sélection exige de longs calculs et n'est possible qu'avec l'aide d'un logiciel de plans d'expériences.

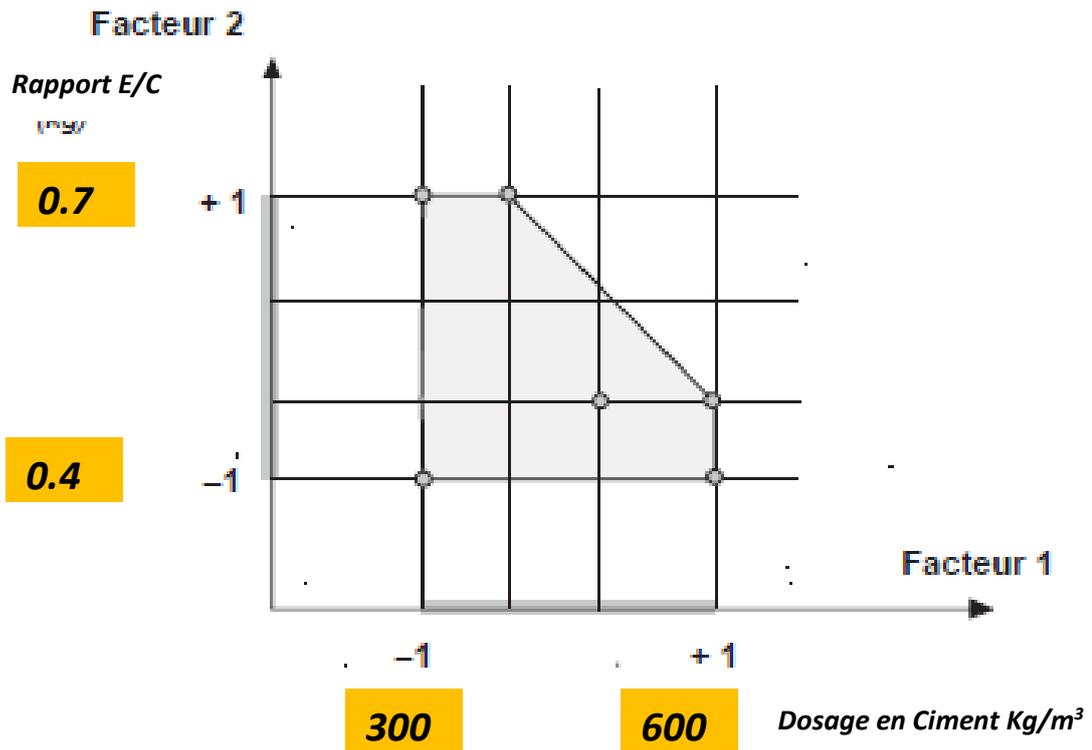


Figure 1.9 – Les meilleurs points sont sélectionnés par le logiciel.

### 1.6.3 Surfaces de réponse

À chaque point du domaine d'étude correspond une réponse. À l'ensemble de tous les points du domaine d'étude correspond un ensemble de réponses qui se localise sur une surface appelée *surface de réponse* (figure 1.10).

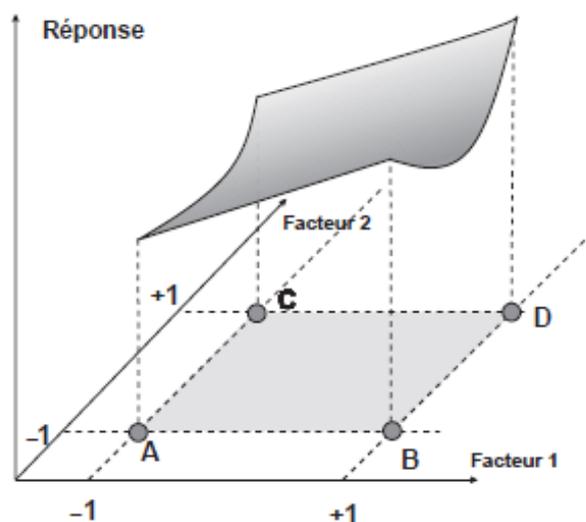


Figure 1.10 – L'ensemble des réponses qui correspond à tous les points du domaine d'étude forme la surface de réponse.

En général, on ne connaît que quelques réponses, celles qui correspondent aux points expérimentaux retenus par l'expérimentateur. On interpole à l'aide d'un modèle mathématique, les réponses inconnues pour obtenir la surface de réponse. Les points d'expériences retenus par la théorie des plans d'expériences assurent la meilleure précision possible sur la forme et la position de la surface de réponse.

#### 1.6.4 Modélisation mathématique a priori de la réponse

- **Modélisation mathématique**

En l'absence de toute information sur la fonction qui lie la réponse aux facteurs, on se donne a priori une loi d'évolution dont la formulation la plus générale est la suivante :

$$y = f(x_1, x_2, x_3, \dots, x_n) \quad \{1.2\}$$

Cette fonction est trop générale et il est d'usage d'en prendre un développement limité de Taylor-Mac Laurin, c'est-à-dire une approximation. Si les dérivées peuvent être considérées comme des constantes, le développement précédent prend la forme d'un polynôme de degré plus ou moins élevé :

$$y = a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j + \sum a_{ii} x_i^2 + \dots \quad \{1.3\}$$

où :

- $y$  est la grandeur à laquelle s'intéresse l'expérimentateur ; c'est la réponse ou la grandeur d'intérêt,
- $x_i$  représente un niveau du facteur  $i$ ,
- $x_j$  représente un niveau du facteur  $j$ ,
- $a_0, a_i, a_{ij}, a_{ii}$  sont les *coefficients* du polynôme.

Ce modèle est appelé le modèle *a priori* ou le modèle *postulé*.

Les modèles établis sont des modèles de prévision valables dans le domaine d'étude, domaine que l'on doit toujours préciser. Ce ne sont pas des modèles théoriques basés sur des lois physico-chimiques ou mécaniques. Dans quelques rares cas, il est possible d'utiliser des lois physiques théoriques connues.

- **Modélisation expérimentale**

Deux compléments doivent être apportés au modèle purement mathématique précédemment décrit.

Le premier complément est le manque d'ajustement. Cette expression traduit le fait que le modèle choisi par l'expérimentateur avant les expériences est probablement un peu différent du modèle réel qui régit le phénomène étudié. Il y a un écart entre ces deux modèles. Cet écart est le *manque d'ajustement* (*lack of fit* en anglais), on le note par la lettre D. Le second complément est la prise en compte de la nature aléatoire de la réponse. En effet, dans le cas général, si l'on mesure plusieurs fois une réponse en un même point expérimental, on n'obtiendra pas exactement le même résultat. Il y a une dispersion des résultats. Les dispersions ainsi constatées sont appelées *erreurs aléatoires* ou *erreurs expérimentales* (*pure error* en anglais) et on les note par la lettre e.

La relation générale {1.2} doit être modifiée ainsi :

$$y = f(x_1, x_2, x_3, \dots, x_n) + D + e \quad \{1.4\}$$

Cette relation sera exploitée au cours du chapitre 5 où l'on verra comment on estime le manque d'ajustement D et l'erreur aléatoire e.

- **Systeme d'equations**

Chaque point expérimental apporte une valeur de la réponse. Or cette réponse est modélisée par un polynôme dont les coefficients sont les inconnues qu'il faut déterminer. À la fin du plan d'expériences, on a un système de  $n$  équations (s'il y a  $n$  essais) à  $p$  inconnues (s'il y a  $p$  coefficients dans le modèle choisi *a priori*). Ce système s'écrit d'une manière simple en notation matricielle :

$$\mathbf{y} = \mathbf{X} \mathbf{a} + \mathbf{e} \quad \{1.5\}$$

où :

- $\mathbf{y}$  est le *vecteur des réponses*,
- $\mathbf{X}$  est la *matrice de calcul des coefficients* ou *matrice du modèle* qui dépend des points expérimentaux choisis pour exécuter le plan et du modèle postulé,
- $\mathbf{a}$  est le *vecteur des coefficients*,
- $\mathbf{e}$  est le *vecteur des écarts*.

Ce système ne peut pas, en général, être résolu simplement car le nombre d'équations est inférieur au nombre d'inconnues. En effet, il y a  $n$  équations et  $p + n$  inconnues. Cette résolution ne peut être menée à bien que si l'on utilise une méthode de régression. La plupart du temps cette méthode est basée sur le critère d'optimisation des moindres carrés. On obtient ainsi les estimations des coefficients que l'on note

$\hat{\mathbf{a}}$

Le résultat de ce calcul est :

$$\hat{\mathbf{a}} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{y} \quad \{1.6\}$$

formule dans laquelle la matrice  $\mathbf{X}'\mathbf{X}$  est la matrice transposée de  $\mathbf{X}$  (voir l'annexe D sur le calcul matriciel). Il existe de nombreux logiciels qui exécutent ce calcul et qui donnent directement les valeurs des coefficients.

Deux matrices interviennent constamment dans la théorie des plans d'expériences :

- la matrice d'information  $\mathbf{X}' \mathbf{X}$ ,
- la matrice de dispersion.  $(\mathbf{X}' \mathbf{X})^{-1}$ .