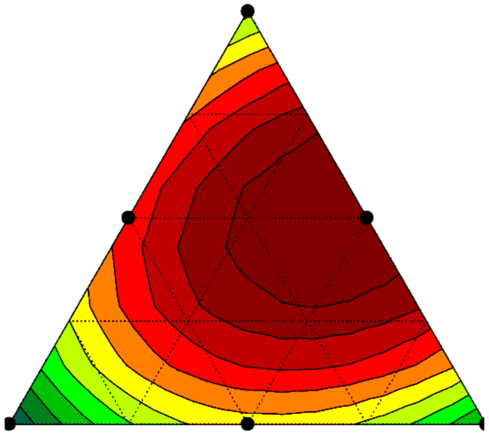
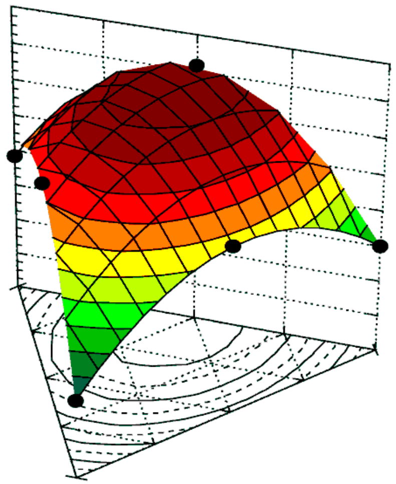
Cours sur les plans de mélange





F. Husson
Laboratoire de mathématiques appliquées
husson@agrocampus-ouest.fr

Transparents largement inspirés de ceux de B. Govaerts

Exemple d’utilisation des plans de mélange

Objectif : une firme alimentaire produit des jus de fruit et désire mettre sur le marché un nouveau cocktail de fruits à base de :

- jus d’orange
- jus de banane
- jus de mangue
- et en ajoutant ou non un additif pour la couleur

- Evaluation du goût par un jury de consommateurs

Qu’est-ce qu’un plan de mélange ?

Mélange :

mixture obtenue en mélangeant différents ingrédients en certaines proportions

Caractéristiques d’un plan de mélange :

• Les facteurs mis en jeu sont les proportions des différents ingrédients (et non les quantités absolues)
• Le domaine des facteurs est contraint : $X_1+X_2+X_3 \ldots + X_k = 1$
• Les réponses ne sont pas influencées par les quantités absolues des facteurs mais par les proportions relatives de ceux-ci

Comment étudier un plan de mélange ?

Modèle pour les plans de mélange :

• Rappel du modèle de régression d'ordre 1 : $Y_i = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ik} + \varepsilon_i$

Mais il faut prendre en compte la contrainte : $X_1 + X_2 + ... + X_k = 1$

nb paramètres

• Modèle d'ordre 1 : $Y_i = \sum_{k=1}^{K} \beta_k x_{ik} + \varepsilon_i$ k

• Modèle d'ordre 2 : $Y_i = \sum_{k=1}^{K} \beta_k x_{ik} + \sum_{k=2}^{K} \sum_{j<k}^{K-1} \delta_{jk} x_{ij} x_{ik} + \varepsilon_i$ k(k+1) / 2

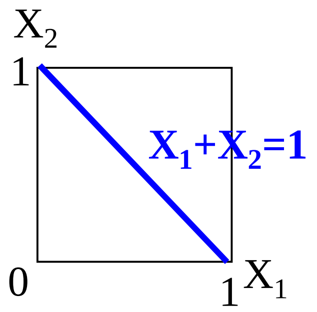
• Modèle d'ordre 3 : prend en compte des interations d'ordre 3 (utilité ?)

Domaine expérimental d'un mélange

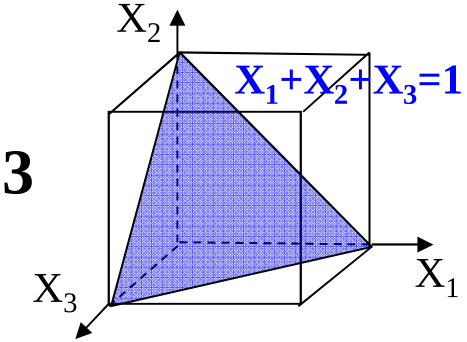
Le domaine expérimental d'un mélange est contraint par la relation :

$$X_1+X_2+X_3 \dots + X_k = 1$$

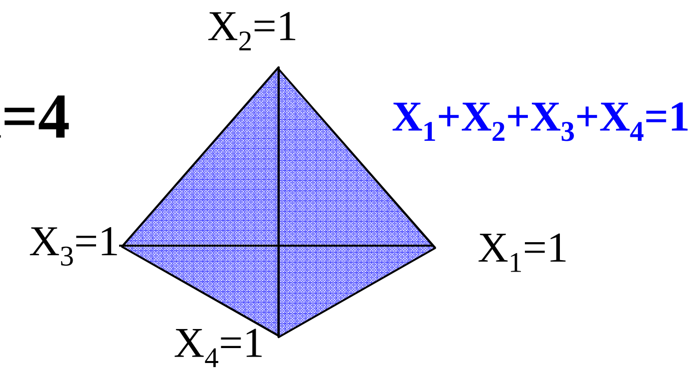
K=2



K=3

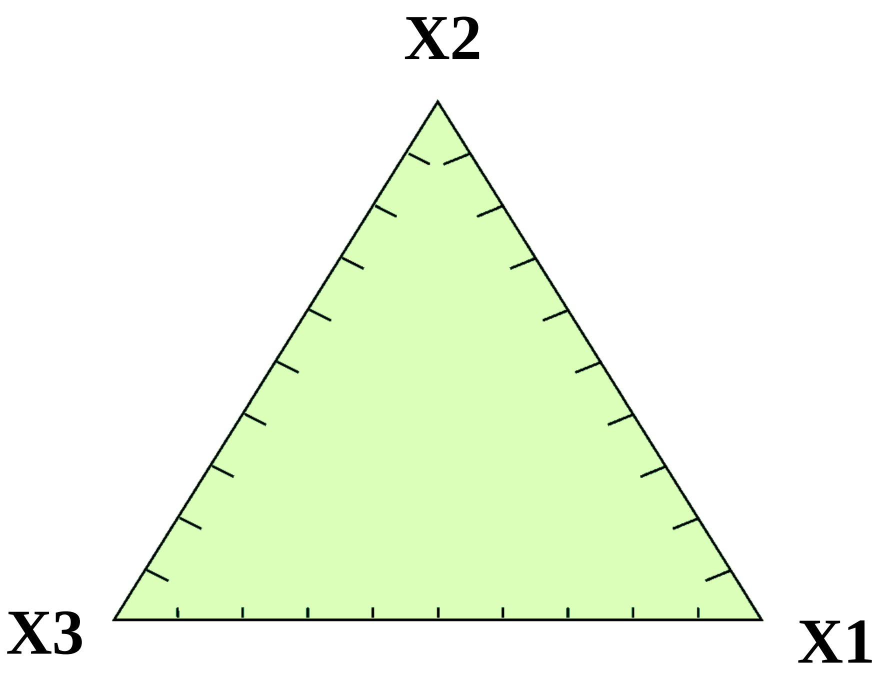


K=4

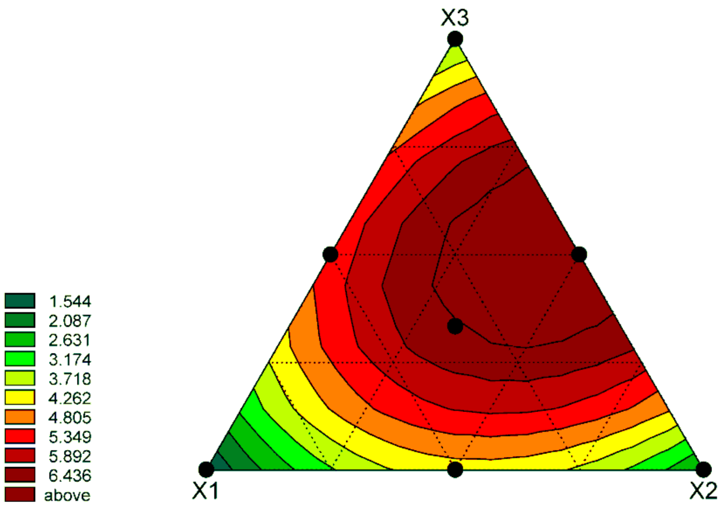


Représentation triangulaire d'un mélange

Représentation triangulaire d'un domaine de mélange à trois facteurs

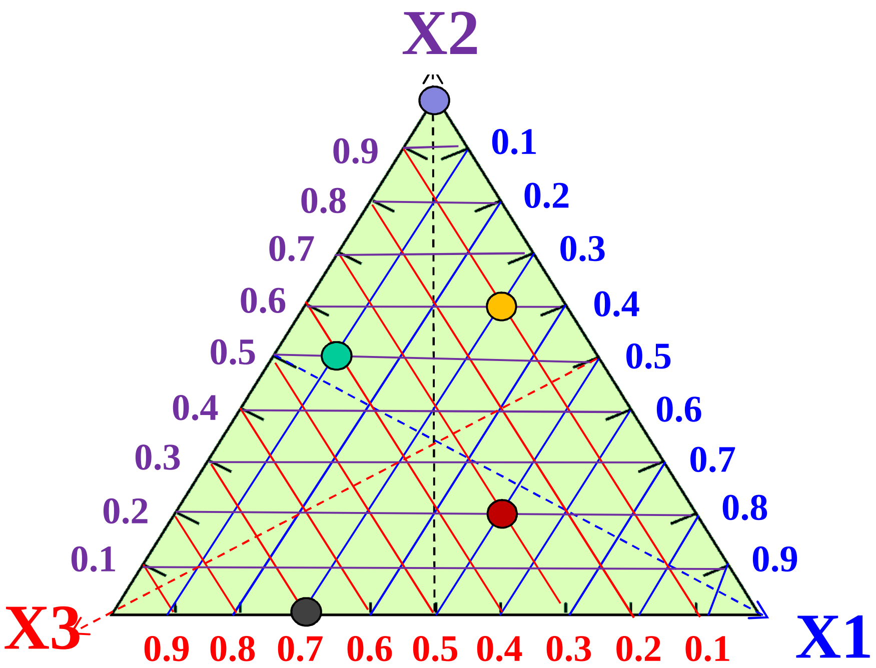


Surface de réponse dans un domaine de mélange à trois facteurs



Représentation triangulaire d'un mélange

Représentation triangulaire d'un domaine de mélange à trois facteurs



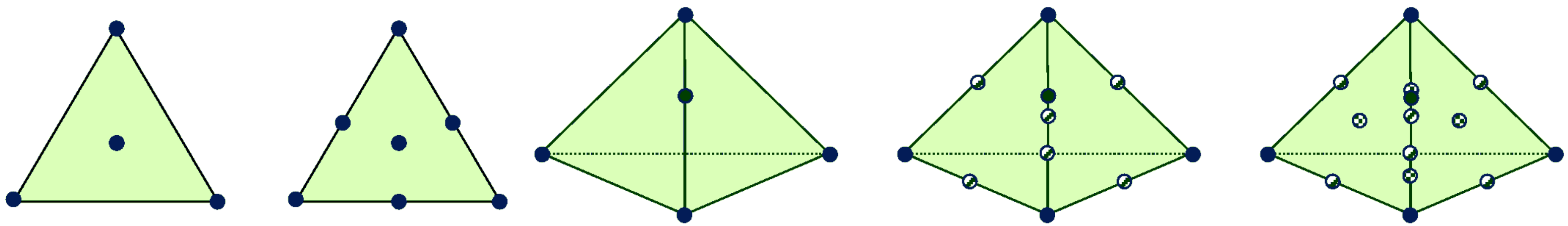
Quel mélange pour les expériences suivantes ?

- (0.1, 0.5, 0.4)
- (0.3, 0, 0.7)
- (0.5, 0.2, 0.3)
- (0.3, 0.6, 0.1)
- (0, 1, 0)

Plan de mélange simplex centroïd

Un plan **simplex centroïd** à k facteurs étudié avec un modèle de degré m est constitué des mélanges suivants ($m < k$) :

- ✓ chaque constituant pur
- ✓ mélange de 2 constituants en proportions égales
- ✓ …
- ✓ mélange de m constituants en proportions égales
- ✓ d'essais au centre du domaine (tous les constituants en proportions égales)

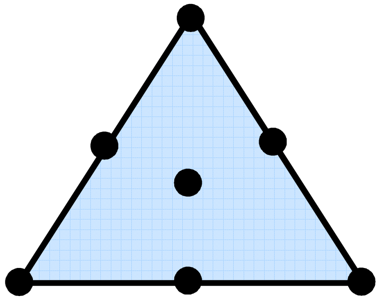


k=3 et m=1 **k=3 et m=2** **k=4 et m=1** **k=4 et m=2** **k=4 et m=3**

Bien souvent en pratique, on se contente de $m \leq 2$

Plan de mélange simplex centroïde

Exemple : plan **simplex centroïd** à 3 facteurs étudié avec un modèle de degré 2

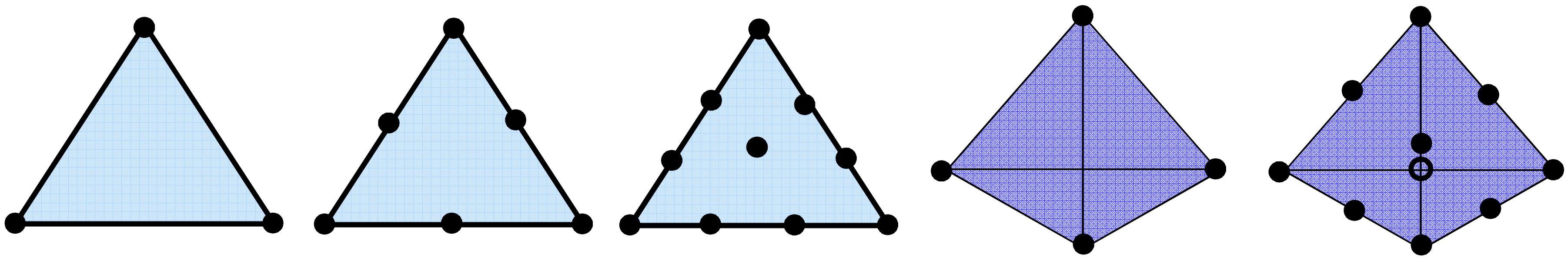


Essai	X1	X2	X3
1	1	0	0
2	0	1	0
3	0	0	1
4	0.5	0.5	0
5	0.5	0	0.5
6	0	0.5	0.5
7	1/3	1/3	1/3

Plan de mélange : réseau simplex (réseau de Scheffé)

Un **réseau simplex** à k facteurs étudié et m+1 niveaux est constitué de toutes les combinaisons possibles de m+1 niveaux pour chaque constituant :

$$(0, 1/m, 2/m, 3/m, \ldots, k/m)$$

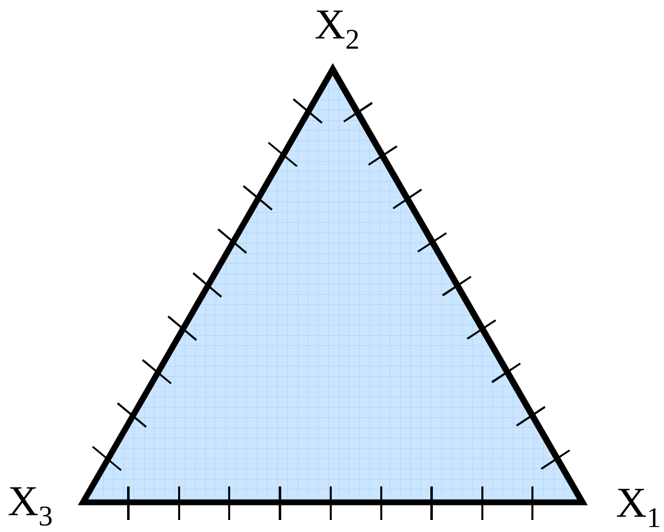


Grille régulière dans le domaine qui peut être considéré comme un plan factoriel complet dans le cas classique des plans

Plan de mélange de type II

Contraintes sur les facteurs : $I_i \leq X_i \leq 1$

Exemple : $0.3 \leq$ orange ≤ 1; $0.2 \leq$ banane ≤ 1; $0.1 \leq$ mangue ≤ 1;



Quelle est la forme du domaine et quel plan utiliser ?

Plan de mélange de type II

Contraintes sur les facteurs : $I_i \leq X_i \leq 1$

Contraintes :

$0.3 \leq$ orange ≤ 1

$0.2 \leq$ banane ≤ 1

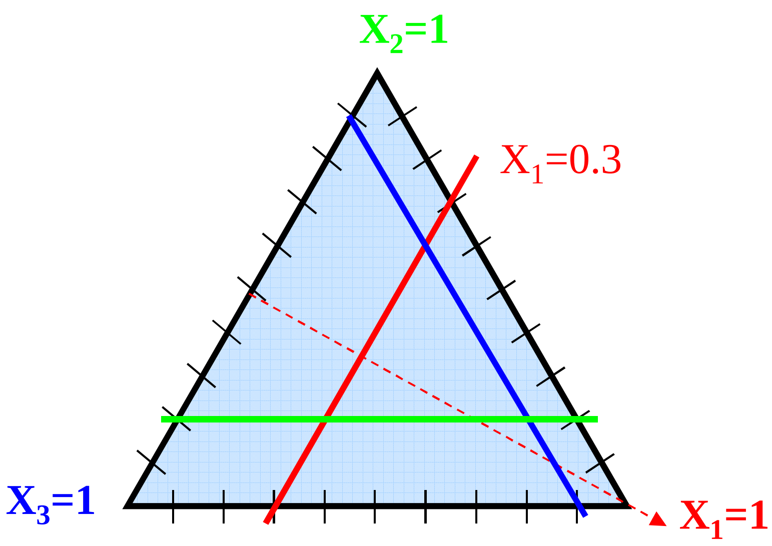
$0.1 \leq$ mangue ≤ 1

Contraintes réelles :

$0.3 \leq$ orange $\leq 1\text{-}0.2\text{-}0.1=0.7$

$0.2 \leq$ banane $\leq 1\text{-}0.3\text{-}0.1=0.6$

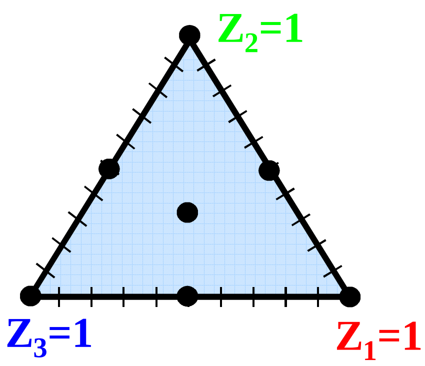
$0.1 \leq$ mangue $\leq 1\text{-}0.3\text{-}0.2=0.5$



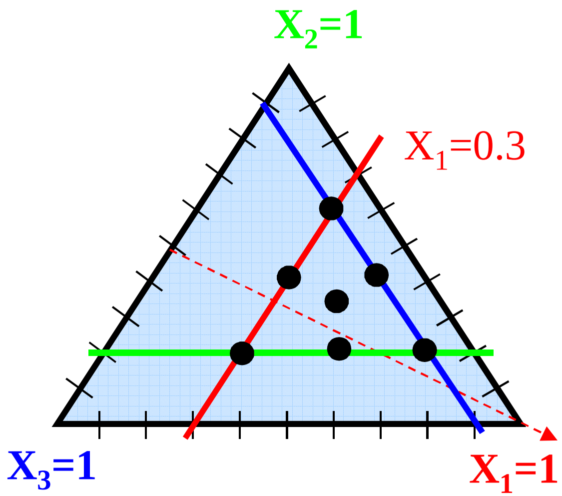
Plan de mélange de type II

Quand les proportions des constituants sont soumises à des contraintes inférieures, le domaine expérimental est un tétraèdre et les plans de mélange classiques peuvent être utilisés après transformation.

Transformation : $X_i = I_i + (1\text{-}I)\, Z_i$ avec $I = I_1 + I_2 + \ldots + I_k$



Z1	Z2	Z3
1	0	0
0	1	0
0	0	1
0.5	0.5	0
0.5	0	0.5
0	0.5	0.5
1/3	1/3	1/3

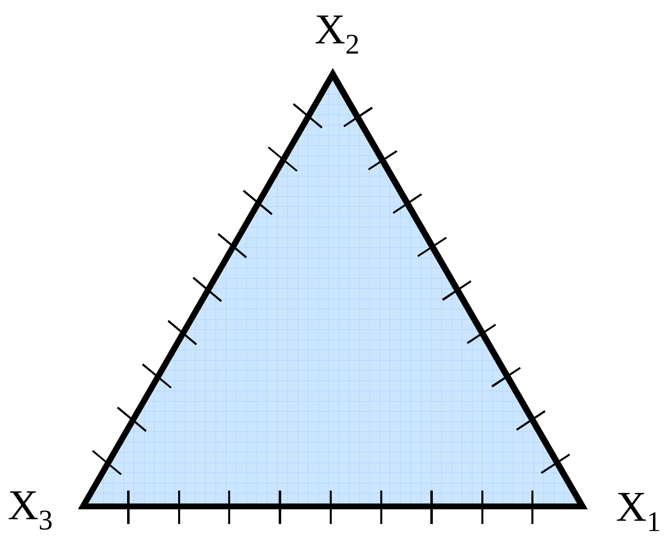


X1	X2	X2
0.7	0.2	0.1
0.3	0.6	0.1
0.3	0.2	0.5
0.5	0.4	0.1
0.5	0.2	0.3
0.3	0.4	0.3
0.43	0.33	0.23

Plan de mélange de type III

Contraintes sur les facteurs : $I_i \leq X_i \leq S_i$

Exemple : $0.3 \leq$ orange ≤ 0.6 ; $0.2 \leq$ banane ≤ 0.5; $0.1 \leq$ mangue ≤ 0.4;



Quelle est la forme du domaine et quel plan utiliser ?

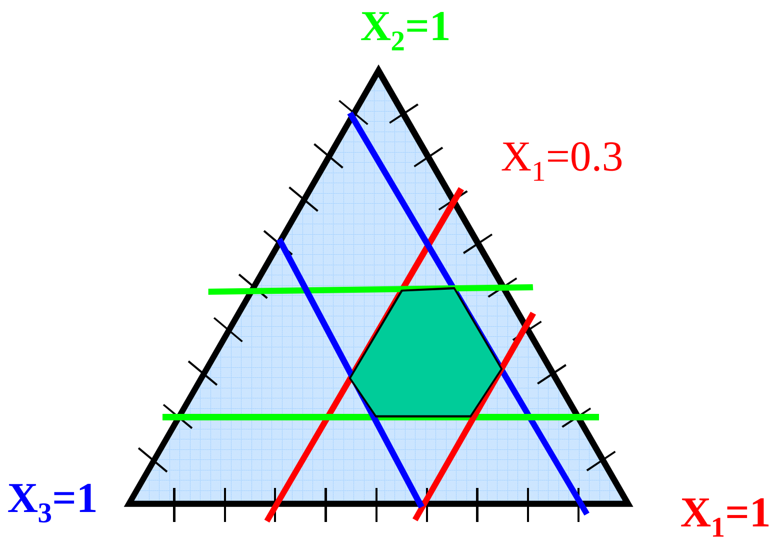
Plan de mélange de type III

Contraintes :

$0.3 \leq$ orange ≤ 0.6

$0.2 \leq$ banane ≤ 0.5

$0.1 \leq$ mangue ≤ 0.4



Quand les proportions des constituants sont soumises à des contraintes inférieures ET supérieures, le domaine expérimental est un polyèdre irrégulier et les plans de mélange classiques ne peuvent plus être utilisés après transformation.

Plan de mélange de type III

Générer un ensemble de *m* essais candidats :

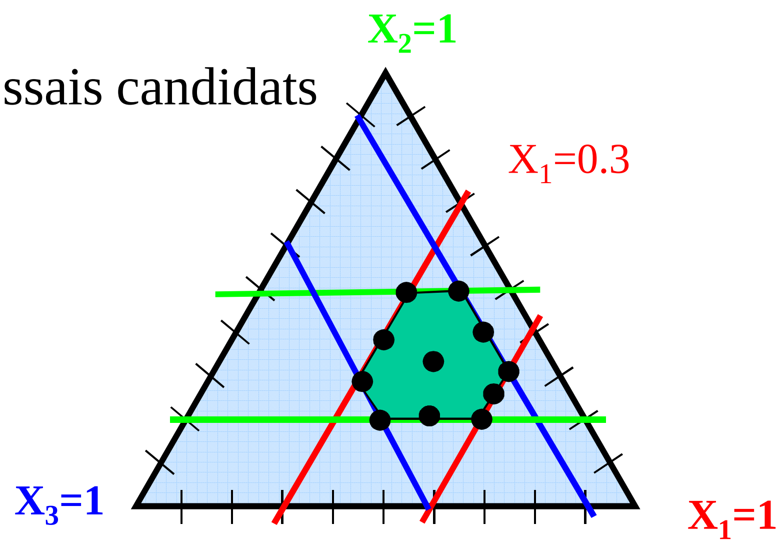
✓Sommets du polyèdre

✓Milieux des arêtes et des faces

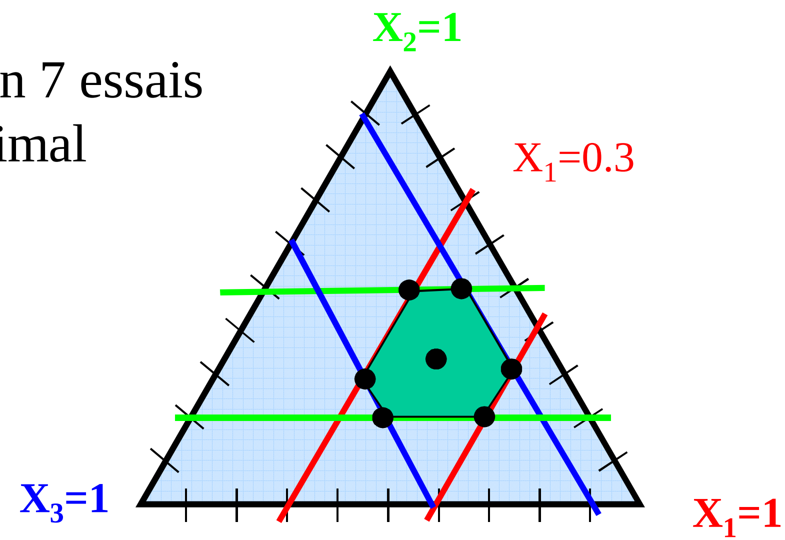
✓Centre du polyèdre

Puis choisir les essais à conserver (algorithme D-optimal)

Essais candidats



Plan en 7 essais
D-optimal



Rappel : qu'est ce qu'un bon plan ?

Un plan tel que l'estimation des effets de chaque variable est la plus précise possible

Il faut minimiser : $V(\widehat{\beta}) = (X'X)^{-1}\sigma^2$

Dépend uniquement du **choix** des expériences

Variabilité résiduelle : dépend des résultats des expériences

→ trouver les expériences telles que $(X'X)^{-1}$ soit minimale

Plan D-optimal minimise le déterminant de la matrice $(X'X)^{-1}$

Plans D-optimaux

Principe :

- Beaucoup d'essais candidats
- N essais choisis au hasard puis algorithme d'échange : échange entre un essai du plan et un essai hors plan effectué si cela diminue le déterminant de $(X'X)^{-1}$
- L'algorithme de Federov s'arrête quand aucun échange n'améliore le critère

Avantages :

- Plan très flexibles par rapport au nombre d'essais
- Possibilité d'imposer des essais
- Démarche séquentielle possible (modèle du 1er ordre, si le modèle n'est pas adapté, l'algorithme fournit les essais supplémentaires)

Inconvénient : un plan est toujours fourni. Est-il de bonne qualité ?

Plan de mélange de type IV et V

Mélange de type IV

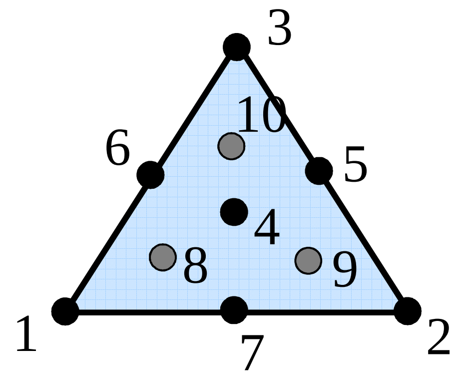
Un des facteurs est très dominant, les autres en faible quantité.
Exemple : Eau entre 99% et 99.5%, acide citrique entre 0.5% et 1%, carraghénane entre 0.5% et 1%
Considérer un plan d'expérience pour facteurs quantitatifs sur les additifs uniquement (et compléter avec l'eau)

Mélange de type V

Problème avec des facteurs de mélange et des facteurs quantitatifs et/ou qualitatifs
Solution : Répéter le plan de mélange pour les différents niveux des facteurs quantitatifs ou qualitatifs
Si cela engendre trop d'essais, extraire un plan D-optimal à partir de cette liste d'essais candidats

Démarche statistique

Construire un plan avec peu d’essais et estimer le modèle avec les effets linéaires.
Valider le modèle grâce au point au centre
si le modèle n’est pas validé, ajouter des essais et considérer
un modèle avec effets quadratiques
Valider le modèle avec les point au centre des domaines



Exemple pour un mélange ternaire :

- Le réseau Scheffé {3;1} comporte 3 points pour un polynôme de degré 1
- Faire les essais pour les points 1, 2 et 3 ce qui permet d’estimer 3 paramètres a1, a2, a3
- Calculer la prévision pour le centre du domaine et faire l’essai de validation,
- Si le modèle du 1er degré est rejeté (à l’incertitude de mesure près), faire l’hypothèse d’un modèle quadratique :
- Le réseau Scheffé {3;2} comporte 6 points pour un polynôme de degré 2,
- Faire les essais supplémentaires 5, 6 et 7 et estimer les 6 paramètres a1, a2, a3, a12, a23, a13
- Calculer les prévisions pour les centres des domaines 8, 9, 10 et éventuellement 4.

Construction de plans avec R

Construction du plan d’expériences centroïd avec 3 ingrédients

```
library(qualityTools )
plan <- mixDesign(3,type="centroid")
  StandOrder RunOrder    Type     A     B     C  y
1          1        6 2-blend 0.000 0.500 0.500 NA
2          2        7 2-blend 0.500 0.500 0.000 NA
3          3        4 2-blend 0.500 0.000 0.500 NA
4          4        1 1-blend 1.000 0.000 0.000 NA
5          5        3    <NA> 0.333 0.333 0.333 NA
6          6        2 1-blend 0.000 0.000 1.000 NA
7          7        5 1-blend 0.000 1.000 0.000 NA
```

Construction d’un réseau de Scheffé

```
plan2 <- mixDesign(3,type="lattice")
   StandOrder RunOrder    Type         A         B         C  y
1           1        9 2-blend 0.0000000 0.3333333 0.6666667 NA
2           2        5 1-blend 0.0000000 0.0000000 1.0000000 NA
3           3        4 2-blend 0.3333333 0.0000000 0.6666667 NA
4           4        7 2-blend 0.6666667 0.0000000 0.3333333 NA
5           5        3 2-blend 0.6666667 0.3333333 0.0000000 NA
6           6        8 2-blend 0.0000000 0.6666667 0.3333333 NA
7           7        6 2-blend 0.3333333 0.6666667 0.0000000 NA
8           8        1 1-blend 0.0000000 1.0000000 0.0000000 NA
9           9        2 1-blend 1.0000000 0.0000000 0.0000000 NA
10         10       10    <NA> 0.3333333 0.3333333 0.3333333 NA
```

Construction de plans avec contraintes sous R

Construction du plan d'expériences centroïd avec contraintes inférieures

```
library(qualityTools )
plan <- mixDesign(3,type="centroid",lower=c(.3,.2,.1))
names(plan@design) = c("orange", "banane", "mangue")
plan@design
```

```
     orange    banane    mangue
1 0.7000000 0.2000000 0.1000000
2 0.5000000 0.4000000 0.1000000
3 0.4333333 0.3333333 0.2333333
4 0.5000000 0.2000000 0.3000000
5 0.3000000 0.6000000 0.1000000
6 0.3000000 0.2000000 0.5000000
7 0.3000000 0.4000000 0.3000000
```

Construction de plans optimaux sous R

Construction d'un plan de mélange avec un facteur quantitatif

```
library(qualityTools )
plan <- mixDesign(3,type="centroid",lower=c(.3,.2,.1))
plan2 <- rbind.data.frame(plan@design,plan@design,plan@design)
names(plan2) <- c("orange", "banane", "mangue")
fac.quanti <- c(rep(5,7),rep(10,7),rep(15,7))
plan3 <- cbind.data.frame(plan2,fac.quanti)
```

contraintes sur les var. de mélange

facteur quantitatif

Construction d'un plan optimal en 12 essais

```
library(AlgDesign)
planD=optFederov(~ -1+(orange+banane+mangue)^2+fac.quanti+I(fac.quanti^2),
                 plan3,nTrials=12)
   orange banane mangue fac.quanti
1     0.3    0.6    0.1          5
2     0.5    0.4    0.1          5
4     0.7    0.2    0.1          5
6     0.3    0.4    0.3          5
9     0.5    0.4    0.1         10
10    0.3    0.2    0.5         10
11    0.7    0.2    0.1         10
12    0.5    0.2    0.3         10
15    0.3    0.6    0.1         15
17    0.3    0.2    0.5         15
19    0.5    0.2    0.3         15
20    0.3    0.4    0.3         15
```

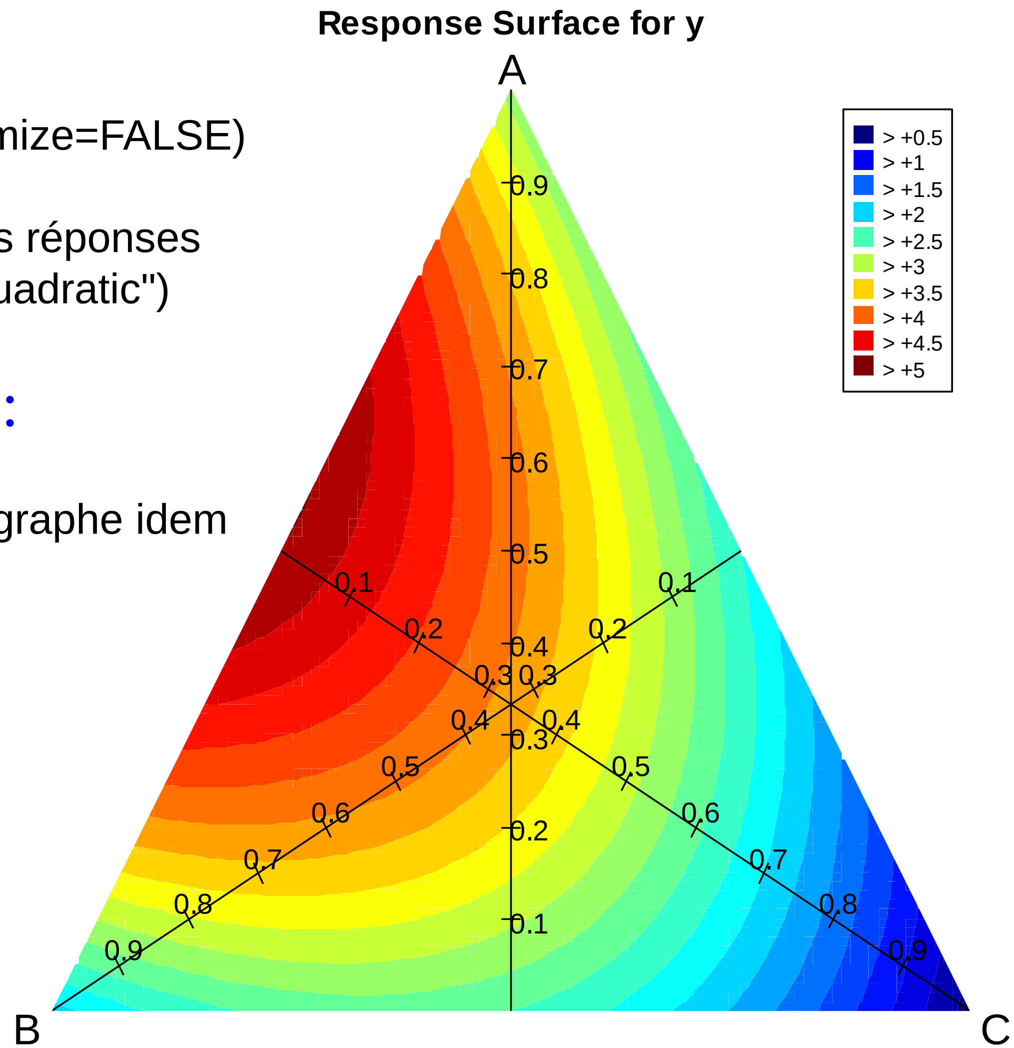
Représentation de surfaces de réponse avec R

Dessin de la surface de réponse (possible si plan obtenu avec la fonction mixDesign)

```
library(qualityTools )
plan <- mixDesign(3,type="centroid",randomize=FALSE)
y <- c(1,2.4,3.2,1.4,1.2,3.6,6.4)
response(plan) <- y                    ## ajout des réponses
contourPlot3(A,B,C,y, data=plan, form = "quadratic")
```

On peut aussi définir le modèle :

```
contourPlot3(A,B,C,y, data=plan,
      form=formula(y~ -1+(A+B+C)^2)) ## graphe idem
```



Détermination d'un optimum

Estimation des paramètres du modèle

```
mod=lm(Gout~ -1+(X1+X2+X3)^2+X4+ I(X4^2),data=don)
summary(mod)
          Estimate Std. Error t value Pr(>|t|)
X1       -19.91189    4.03501  -4.935 0.001685 **
X2       -17.96241    4.10318  -4.378 0.003244 **
X3       -11.54983    5.09151  -2.268 0.057609 .
X4         1.04690    0.13971   7.493 0.000138 ***
I(X4^2)   -0.05935    0.00708  -8.382 6.76e-05 ***
X1:X2     67.91227   14.70224   4.619 0.002429 **
X1:X3     52.47022   14.65999   3.579 0.008986 **
X2:X3     43.43361   11.18708   3.882 0.006034 **
```

Données

X1	X2	X3	X4	Gout
0.4	0.5	0.1	5	3.25
0.4	0.5	0.1	5	3.75
0.55	0.2	0.25	5	3
0.3	0.3	0.4	5	5.25
0.3	0.3	0.4	5	5.75
0.42	0.34	0.24	5	5
0.55	0.35	0.1	10	3.5
0.3	0.5	0.2	10	4.75
0.3	0.5	0.2	10	5.25
0.4	0.2	0.4	10	5.5
0.42	0.34	0.24	10	6
0.55	0.35	0.1	15	1.5
0.55	0.2	0.25	15	2
0.4	0.2	0.4	15	3.5
0.3	0.4	0.3	15	3.5

Détermination de l'optimum avec la fonction optim

```
fct.a.optim = function(x){
  x1=x[1]   ;  x2=x[2]   ; x4=x[3]
  x3=1-x[1]-x[2] ### On impose la contrainte sur un des paramètres
  Y = -19.9*x1-18*x2-11.5*x3+1.05*x4+67.9*x1*x2+52.47*x1*x3+43.4*x2*x3-0.059*x4^2
  Y = -Y  ### La fonction optim trouvant un minimum, on minimise l'opposé de Y
}
yy=optim(c(0.3,0.3,10),fct.a.optim,lower=c(0,0,5),upper=c(1,1,15))
```

initialisation des paramètres

Bornes inf des paramètres

Bornes sup des paramètres

```
$par
[1] 0.3354582 0.3170189 8.8983053
$value
[1] -6.412429
```

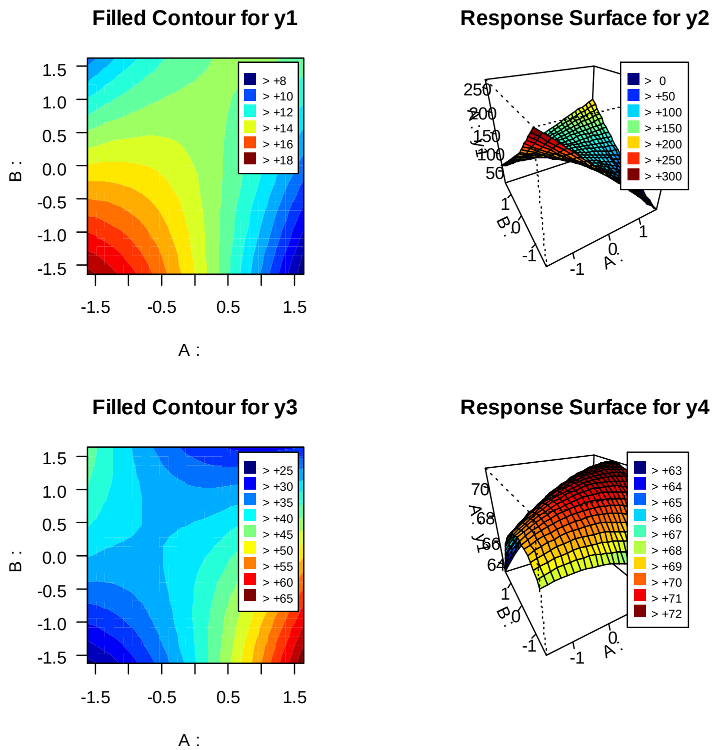
Détermination d'un optimum sur plusieurs réponses

Attention, fonctionne pour plans factoriels (pas pour les mélanges)

```
# Construire le plan
library(qualityTools)
plan = rsmDesign(k = 3, alpha = 1.633, cc = 0, cs = 6)

# Attacher les 4 réponses
y1 = c(10,12,12,20,10,13,13,14,10,15,9,16,12,15,13,13,14,14,15,14)
y2 = c(90,86,80,229,49,129,127,109,77,169,70,154,218,178,130,130,114,109,126,134)
y3 = c(47,41,57,24,64,27,41,38,59,26,52,38,52,29,38,38,43,43,39,39)
y4 = c(67,65,77,75,62,67,78,70,76,70,63,75,65,71,70,68,68,68,69,70)
response(plan) = data.frame(y1, y2, y3, y4)[c(5,2,3,8,1,6,7,4,9:20),]

# Définir les noms et les valeurs réelles des facteurs
lows(plan) = c(0.7, 40, 1.8)
highs(plan) = c(1.7, 60, 2.8)

# Fixer les valeurs acceptables par variable
desires(plan) = desirability(y1, 12, 17, target = "max")
desires(plan) = desirability(y2, 100, 130, target = "max")
desires(plan) = desirability(y3, 40, 60, target = 50)
desires(plan) = desirability(y4, 60, 75, target = 67.5)
desires(plan)
```

Détermination d’un optimum sur plusieurs réponses

Attention, cela ne fonctionne pas pour des mélanges pour l’instant, mais uniquement pour des plans factoriels




```
# Voir les surfaces de réponses
par(mfrow = c(2,2))
contourPlot(A, B, y1, data = plan)
wirePlot(A, B, y2, data = plan)
contourPlot(A, B, y3, data = plan)
wirePlot(A, B, y4, data = plan)
```

```
# Choisir les modèles
fits(plan) = lm(y1 ~ A + B + C + A:B + A:C + B:C + I(A^2) + I(B^2) + I(C^2), data = plan)
fits(plan) = lm(y2 ~ A + B + C + A:B + A:C + B:C + I(A^2) + I(B^2) + I(C^2), data = plan)
fits(plan) = lm(y3 ~ A + B + C + A:B + A:C + B:C + I(A^2) + I(B^2) + I(C^2), data = plan)
fits(plan) = lm(y4 ~ A + B + C + A:B + A:C + B:C + I(A^2) + I(B^2) + I(C^2), data = plan)
```

```
# Trouver l’optimum
## optimum(plan, type = "optim")  ## tps de calcul très long
optimum(plan, type = "grid")         ## bonne approx de l'optimum
```

Construction de plans avec R
package RcmdrPlugin.DoE

Package très complet MAIS il faut bien connaître les plans d'expériences pour pouvoir l'utiliser

Certaines procédures ne sont pas encore programmées, notamment avec les plans de mélange (construction + dépouillement)