KU LEUVEN

Designs in constrained experimental regions

Design and Analysis of Mixture Experiments

Prof Eddie Schrevens
Biosystems Department
Faculty of Bioscience Engineering
K. Universiteit Leuven
Belgium
Eddie.schrevens@biw.kuleuven.be

Part II

Irregular constrained mixture spaces

KU LEUVEN

Mixture experimental design

- Transform to independent factors
  - Classical design and analysis
- Whole simplex experiments
  - {q,m} simplex lattice, simplex centroid (D-optimal with respect to corresponding mixture models)
- Homomorphic experimental regions
  - Same shape as the overall simplex
  - Pseudo-components transformation expands the experimental region to the whole simplex
  - Whole simplex experiments in pseudocomponents
- Complex constrained experimental regions
  - Irregular, convex hyper-polyhedrons

Complex constrained experimental regions

- Each specific problem has its own optimal design related to size, shape and location of the experimental region defined by the set of additional constraints
- Optimal design theory is indispensable
- Implies computer aided design of experiments, especially in high dimensionality
- Extremely computational intensive
- Ill conditioning is more rule than exception

Complex constrained experimental regions
Classical procedure

- Constraints $\Leftarrow$ expert or screening
- Consistency check of constraints!
- Vertices are computed
- ‘Mixture’ model is chosen
- Centroids of different dimensionality are calculated

↓

List of candidate points: vertices extended with centroids

Complex constrained experimental regions
Classical procedure

- Optimality criteria is chosen: D, G, V, A, …
- Optimal design is selected by an exchange algorithm taking into account the selected model
  - Branch and bound, excursion algorithms or all possible permutations
- This algorithm is run for different numbers of design points

↓

Optimal design for the pre-specified model and optimality criterion

Classical design procedures: example

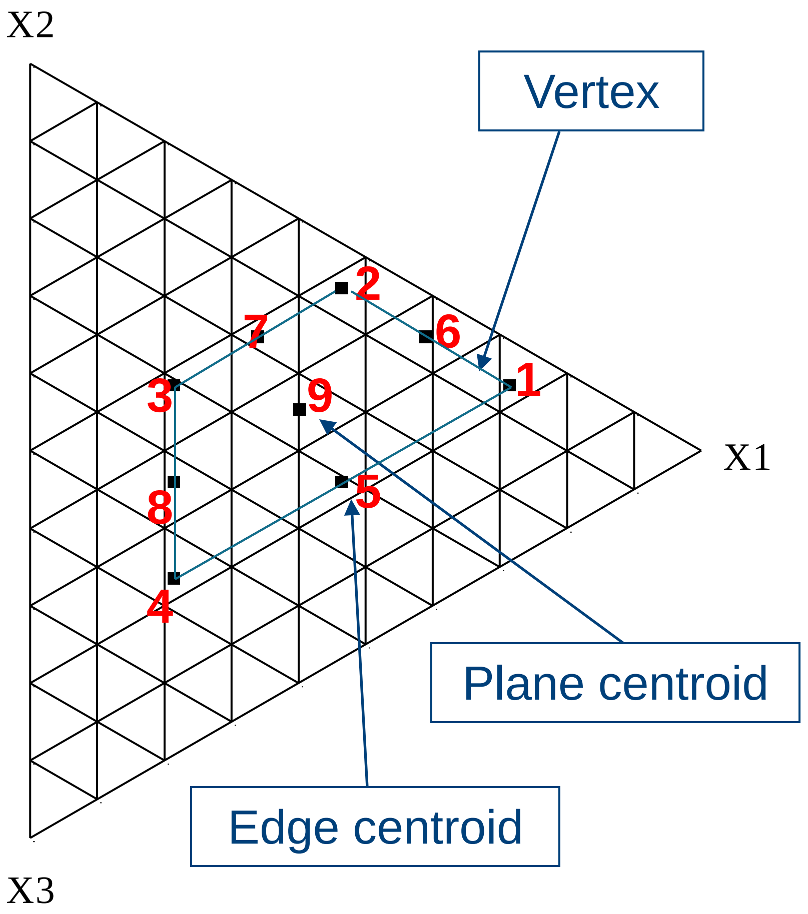
Constraints

$$0.22 \le x_1 \le 0.72$$

$$0.22 \le x_2 \le 0.47$$

$$0.06 \le x_3 \le 0.56$$

List of 9 candidate points



List of candidate points for the linear mixture model

Id	X_1	X_2	X_3
1	0.720	0.220	0.060
2	0.470	0.470	0.060
3	0.220	0.470	0.310
4	0.220	0.220	0.560
5	0.470	0.220	0.310
6	0.595	0.345	0.060
7	0.345	0.470	0.185
8	0.220	0.345	0.435
9	0.408	0.345	0.248

List of candidate points for the quadratic mixture model

Id	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3
1	0.720	0.220	0.060	0.158400	0.043200	0.013200
2	0.470	0.470	0.060	0.220900	0.028200	0.028200
3	0.220	0.470	0.310	0.103400	0.068200	0.145700
4	0.220	0.220	0.560	0.048400	0.123200	0.123200
5	0.470	0.220	0.310	0.103400	0.145700	0.068200
6	0.595	0.345	0.060	0.205275	0.035700	0.020700
7	0.345	0.470	0.185	0.162150	0.063825	0.086950
8	0.220	0.345	0.435	0.075900	0.095700	0.150075
9	0.408	0.345	0.248	0.140760	0.101184	0.085560

Design procedures: example

- Given the expanded design matrix X for OLS estimation, representing model specifications
- For optimal parameter estimation:
  - A-optimality minimises the trace of $(X'X/n)^{-1}$, resulting in a minimal variance of the parameters
  - D-optimality minimizes the determinant of $(X'X/n)^{-1}$
  - E-optimality minimizes the maximum eigenvalue of $(X'X/n)^{-1}$
- For optimal response estimation:
  - G-optimality minimises the maximum prediction variance over a specified set of design points (normalised for the number of design points)
  - V-optimality minimises the average prediction variance over a specified set of design points (normalised for the number of design points)

Design procedures: example, 1st order

- Calculation of optimality criteria for all subsets of 4, 5, 6, 7, 8, 9 points out of the 9 candidates
- Set of points 1,2,3,4 has the lowest D-criterion for a first order linear mixture model

Number of design points	D-optimality	Candidate point number
4	1638	1 2 3 4
5	2133	1 2 3 4 5 1 2 3 4 7
6	2457	1 2 3 4 5 7 1 2 3 4 5 8
7	2804	1 2 3 4 5 6 7 1 2 3 4 5 7 8
8	3013	1 2 3 4 5 6 7 8
9	3812	1 2 3 4 5 6 7 8 9

→ Optimal design

Design procedures: example 2d order

- Calculation of optimality criteria for all subsets of 6, 7, 8, 9 points out of the 9 candidates of the expanded matrix
- Quadratic mixture model
- Set of points 1,2,3,4,6,8 has the lowest D-criterion for a second order linear mixture model

Number design points	D-optimality	Candidate point number
6	3.5 E14	1 2 3 4 6 8
7	5.9 E14	1 2 3 4 5 6 8
8	6.6 E14	1 2 3 4 5 6 8 9
9	8.5 E14	1 2 3 4 5 6 7 8 9

→ Optimal design (row 6: 1 2 3 4 6 8)

How to deal with ill conditioning (multicollinearity) in experiments with mixtures?

- In many situations the additional constraints induce ill-conditioning or near-collinearities in the design matrix on top of the mixture constraint (exact collinearity)
- On top of this, the ill conditioning is increased by model formulation
- Implications for OLS
  - Large estimates
  - Large variances and covariances of the estimates
  - Incorrect signs of the estimates
  - Unreliable test statistics
  - Unreliable variable selection

KU LEUVEN

How to deal with ill conditioning (multicollinearity) in experiments with mixtures?

- Duality
  - Parameter ⇔ predicted response estimation
- Investigate ill conditioning of a selected optimal design
- Decrease ill conditioning by re-design
- Decrease ill conditioning by re-modelling

Investigate collinearity structures
Classical approaches: VIF

OLS of $y=X\beta$ and svd of $X=ULV'$

$$\mathrm{var}(\beta) = \sigma^2 (X'X)^{-1} = \sigma^2 (VL^{-2}V')$$

For the k^{th} element of β

$$\mathrm{var}(\beta_k) = \sigma^2 \sum_{j=1}^{q} \frac{v_{kj}^2}{l_j^2} = \sigma^2 VIF_k$$

$$VIF = \mathrm{diag}\,(X'X)^{-1} = \mathrm{diag}(V'L^{-2}V)$$

Investigate collinearity structures of constrained mixture design matrices: Classical approaches: VIF

$$w_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{n} x_{ij}^{2}}}$$

Standardizing the terms of the model generates the most stable parameter estimation

$W = ULV'$

$VIF = diag\,(W'W)^{-1} = diag(V'L^{-2}V)$

Investigate collinearity structures of constrained mixture design matrices Classical approaches: VIF

$$VIF_i = \sum_{j=1}^{q} \frac{v_{ij}^2}{l_j^2}$$

With X or W = ULV'

VIF_i can be decomposed in components associated with the respective singular values

VIF decomposition

$$VIF_{ij} = \frac{v_{ij}^2}{l_j^2}$$

The proportion of VIF_i Corresponding to l_j

$$\Pi(VIF_i) = \frac{VIF_{ij}}{VIF_i}$$

High variance inflation can be allocated to eigenvectors (PC) with small corresponding singular values l_j

KU LEUVEN

Investigate collinearity structures of constrained mixture design matrices

- Different factorisation of the standardized design matrix
  - Factorisation 1: correlational structure
  - Factorisation 2: collinearity constraints

- Graphical representation by Biplots
  - Rank 2 approximation

Factorisation 1: correlational structure

Starting from the standardized design matrix Z , the correlation matrix $R=Z'Z$

Singular value decomposition: $Z=ULV'$

Factorisation 1 $Z=GH'$

With	$G=U$ and $H=VL$	
Then	$R=H'H$	Correlation
	$ZR^{-1}Z'=GG'$	Mahalanobis distances
	G	Principal components scores

Superimposing the first and second columns of G and H in a biplot results in a rank 2 graphical approximation of the correlational structure of the design

Factorisation 1: correlational structure

Goodness of fit statistics for

R: Correlation matrix

$$GFS(R) = \frac{l_1^4 + l_2^4}{\sum_{i=1}^{q} l_i^4}$$

Z: Standardized design matrix

$$GFS(ZR^{-1}Z) = \frac{l_1 + l_2}{\sum_{i=1}^{q} l_i}$$

ZR^{-1}Z’: Mahalanobis distances

$$GFS(Z) = \frac{l_1^2 + l_2^2}{\sum_{i=1}^{q} l_i^2}$$

Factorisation 2: collinearity constraints

Starting from the standardized design matrix Z

Singular value decomposition: $Z=ULV'$

Factorisation 2 $Z=PQ'$

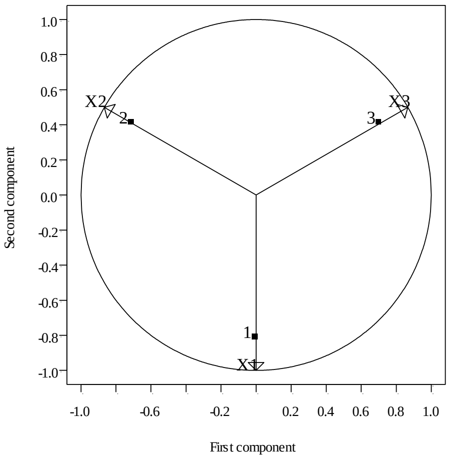
With $P=UL$ and $Q=V$

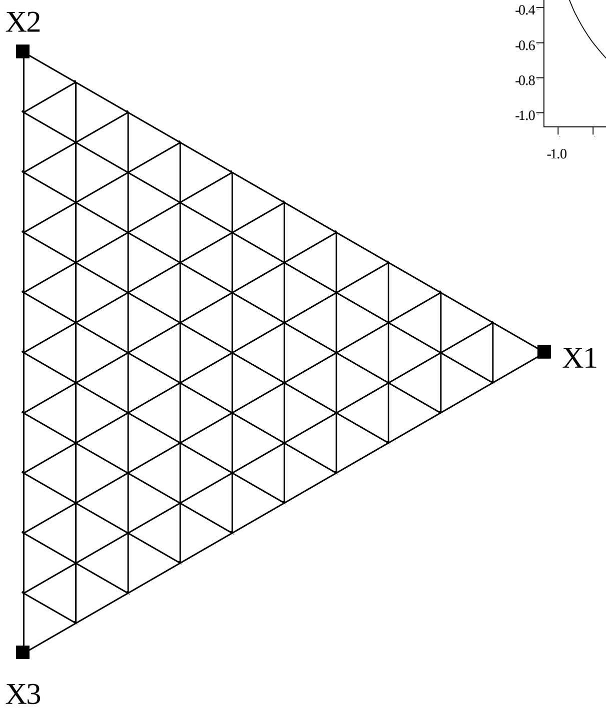
Then Q: principal components
P: principal components scores

Superimposing the columns of P and Q, corresponding to the smallest singular values results in a rank 2 graphical approximation or biplot, that gives information about collinearity equations

Factorisation 1: correlational structure
Example: {3,1} simplex lattice

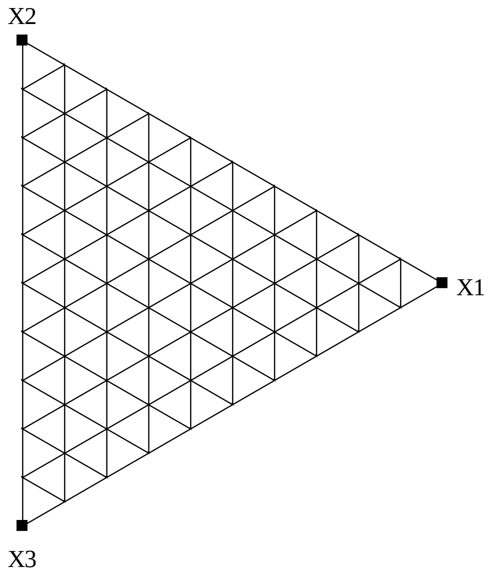
Id	X_1	X_2	X_3
1	1	0	0
2	0	1	0
3	0	0	1





Factorisation 1: correlational structure
Example: {3,1} simplex lattice

Id	X1	X2	X3
1	1	0	0
2	0	1	0
3	0	0	1



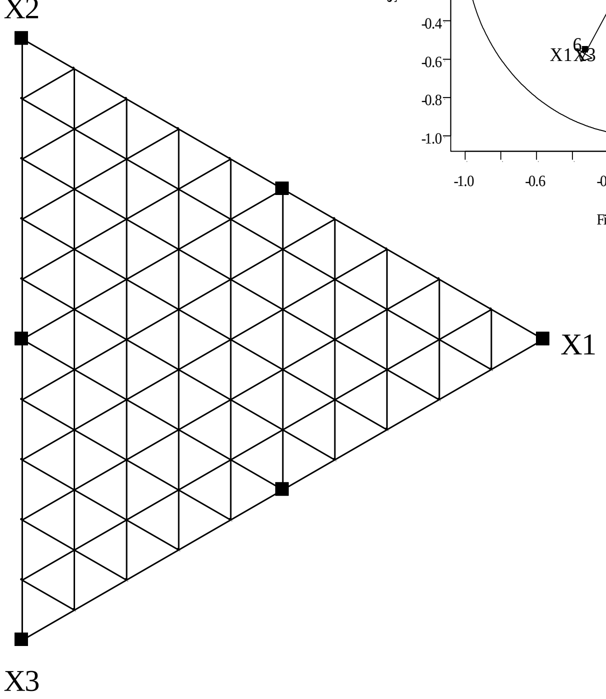
Correlation matrix R		
1	-0.5	-0.5
-0.5	1	-0.5
-0.5	-0.5	1

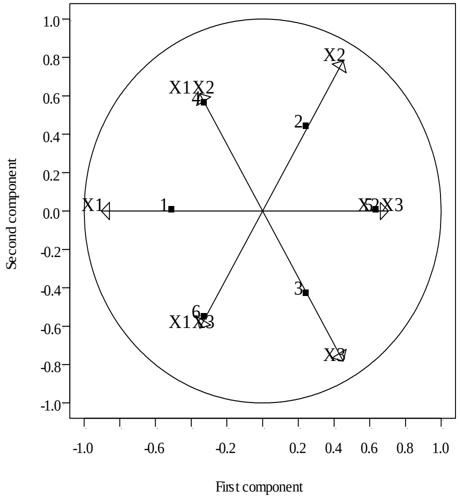
Singular values of X
1
1
1

VIF decomposition			
VIF	VIF_1	VIF_2	VIF_3
1	1	0	0
1	0	1	0
1	0	0	1

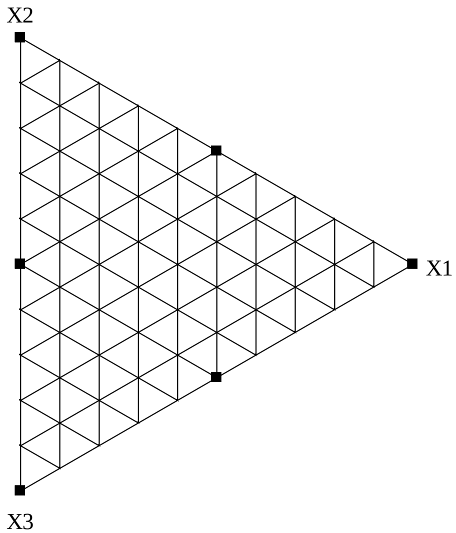
Factorisation 1: correlational structure
Example: {3,2} simplex lattice

Id	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3
1	1.0	0.0	0.0	0.00	0.00	0.00
2	0.0	1.0	0.0	0.00	0.00	0.00
3	0.0	0.0	1.0	0.00	0.00	0.00
4	0.5	0.5	0.0	0.25	0.00	0.00
5	0.0	0.5	0.5	0.00	0.00	0.25
6	0.5	0.0	0.5	0.00	0.25	0.00





Factorisation 1: correlational structure
Example: {3,2} simplex lattice



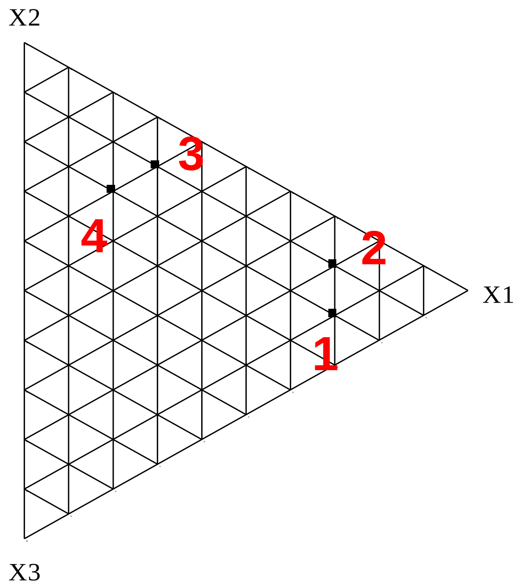
Correlation matrix					
1	-0.5	-0.5	0.2	0.2	-0.4
-0.5	1	-0.5	0.2	-0.4	0.2
-0.5	-0.5	1	-0.4	0.2	0.2
0.2	0.2	-0.4	1	-0.2	-0.2
0.2	-0.4	0.2	-0.2	1	-0.2
-0.4	0.2	0.2	-0.2	-0.2	1

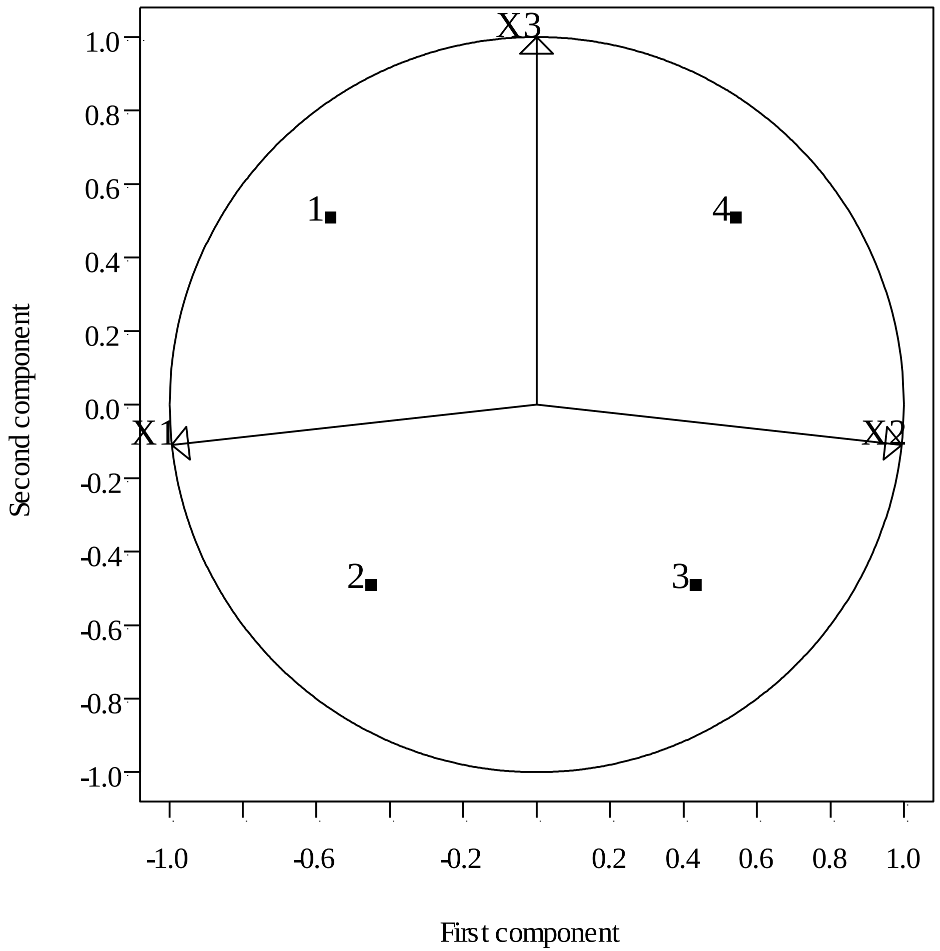
Singular values of W
1.41
1.15
1.15
0.7
0.7
0.5

VIF decomposition							
	VIF	VIF_1	VIF_2	VIF_3	VIF_4	VIF_5	VIF_6
β_1	1.5	0.1	0.2	0	0	0.8	0.4
β_2	1.5	0.1	0.05	0.15	0.6	0.2	0.4
β_3	1.5	0.1	0.05	0.15	0.6	0.2	0.4
β_{12}	1.5	0.066	0.07	0.22	0.4	0.13	0.6
β_{13}	1.5	0.066	0.07	0.22	0.4	0.13	0.6
β_{23}	1.5	0.066	0.3	0	0	0.53	0.6

Factorisation 1: correlational structure
Example: constrained mixture

Id	X_1	X_2	X_3
1	0.7	0.1	0.2
2	0.7	0.2	0.1
3	0.3	0.6	0.1
4	0.2	0.6	0.2

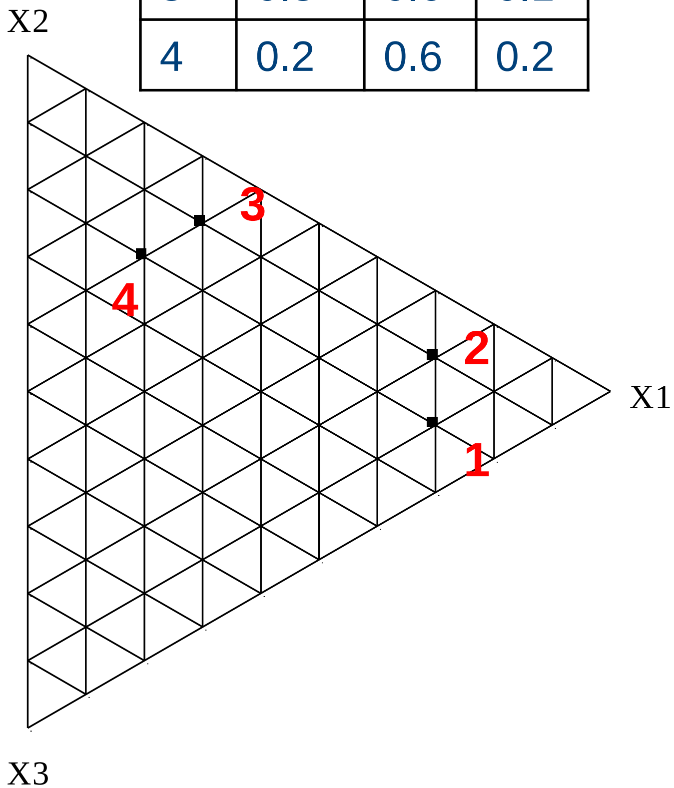




KU LEUVEN

Factorisation 1: correlational structure
Example: constrained mixture

Id	X_1	X_2	X_3
1	0.7	0.1	0.2
2	0.7	0.2	0.1
3	0.3	0.6	0.1
4	0.2	0.6	0.2



Correlation matrix		
1	-0.97	-0.11
-0.97	1	-0.11
-0.11	-0.11	1

Singular values of W
1.57
0.67
0.29

		VIF decomposition		
	VIF	VIF_1	VIF_2	VIF_3
β_1	3.87	0.1283436	0.9900934	2.7530263
β_2	3.06	0.1217495	1.2260788	1.7133913
β_3	7.25	0.1558498	0.005732	7.0896378

KU LEUVEN

Factorisation 1: correlational structure

Example: optimization of bread dough composition

Constraints

Component	Lower bounds	Upper bounds
Water	0.2	0.4
Flour	0.5	0.8
Salt	0.03	0.044
Additive	0.0091	0.0095
Yeast	0.0045	0.0048

→ 16 vertices

↓

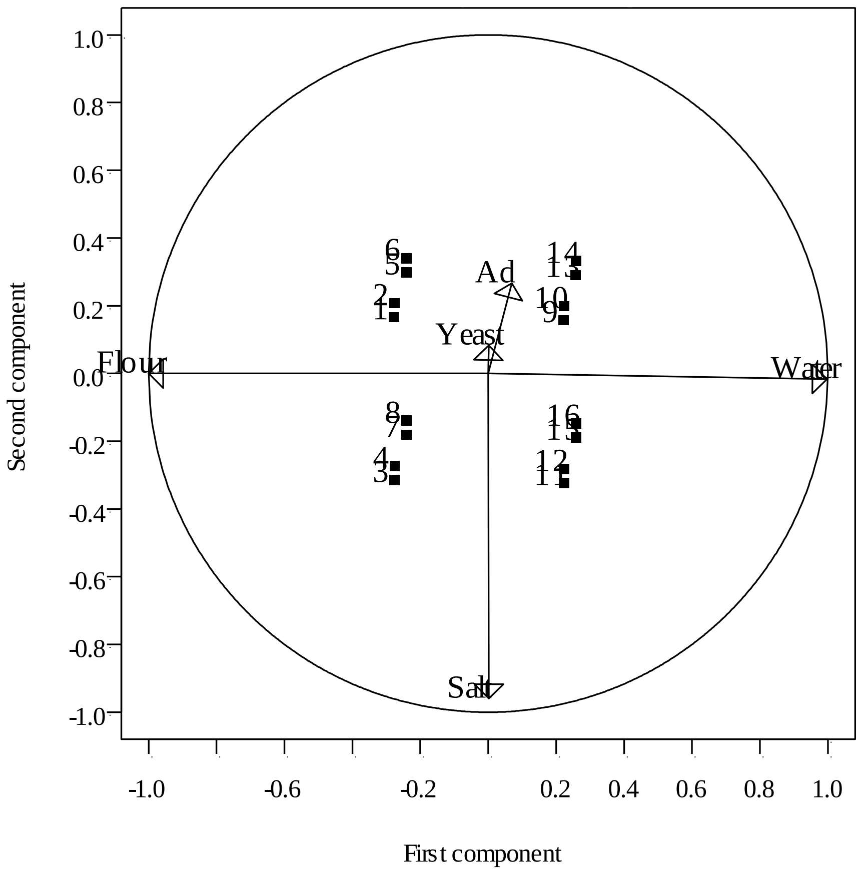
Extended with 65 2,3,4,5 dimensional centroids

↓

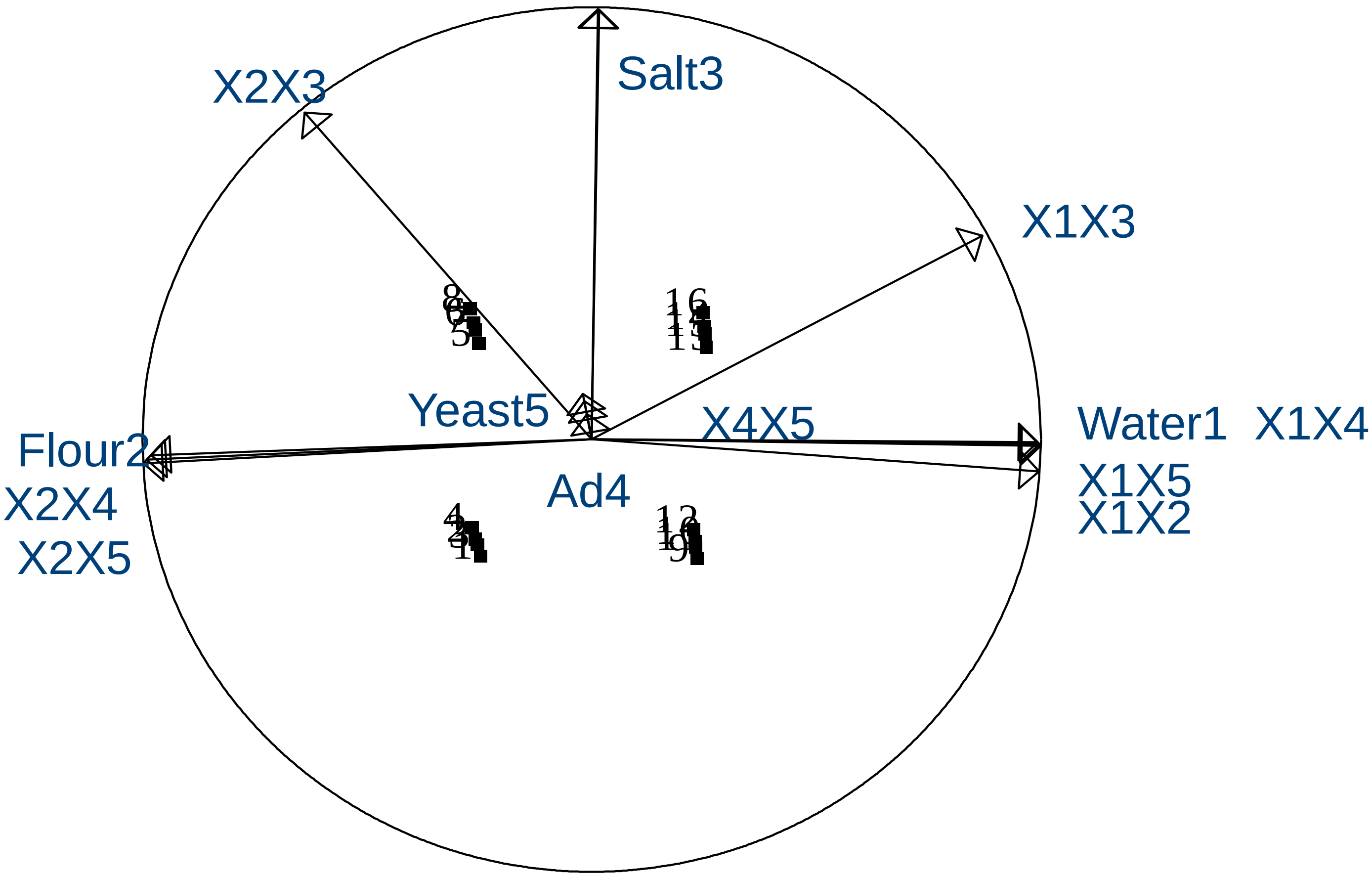
81 candidate points

→ 16 points D-optimal design

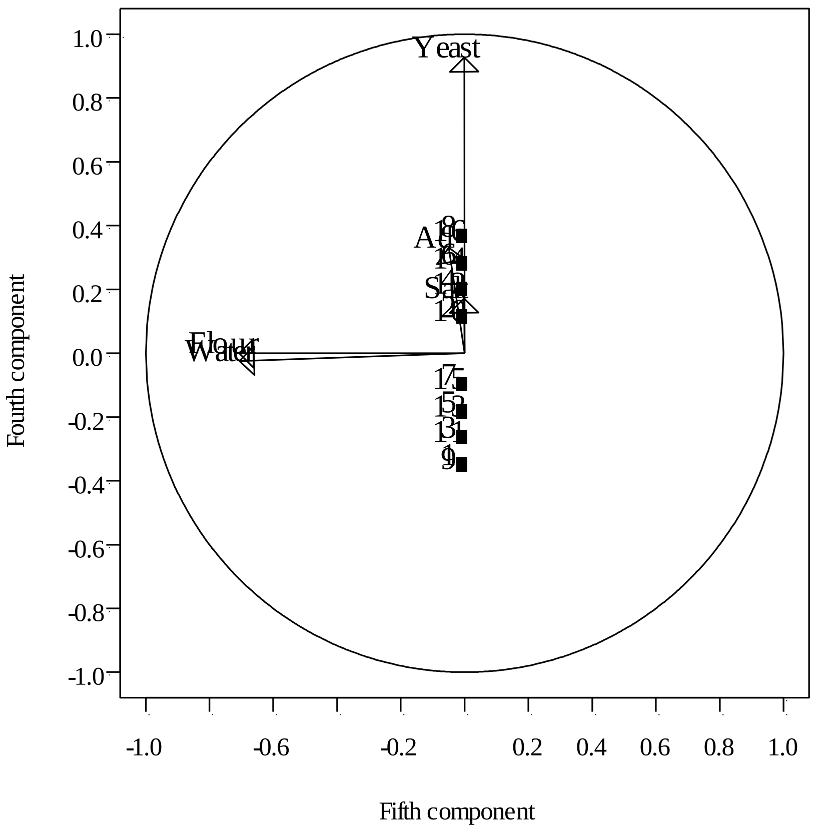
Factorisation 1: correlational structure



Factorisation 1: correlational structure



Factorisation 2: collinearity constraints



Singular values of Z
1.41
1.00
1.00
1.00
2.120 E-16

PC_5 = -0.7 Water -0.7 Flour - 0.005 Salt -0.001Ad - 0.001Yeast $\cong$ 0

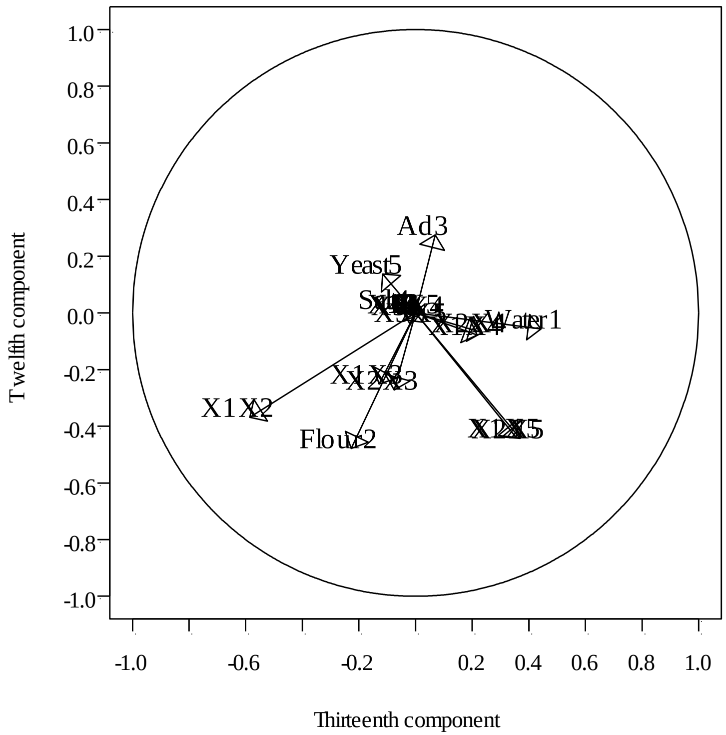
'Near' collinearity constraint

Factorisation 2: collinearity constraints

		VIF decomposition				
	VIF	VIF1	VIF2	VIF3	VIF4	VIF5
β_1	320.11888	0.0389517	5.8668324	0.5888333	18.047538	295.57672
β_2	1363.468	0.0409593	2.334777	3.4602724	78.632188	1278.9998
β_3	31.279872	0.0414719	0.0990308	27.645318	0.1196886	3.3743626
β_4	2123.2871	0.0423696	0.0644128	1.7034599	254.33501	1867.1418
β_5	953.12182	0.0423552	0.0650038	1.774838	846.05622	105.18341

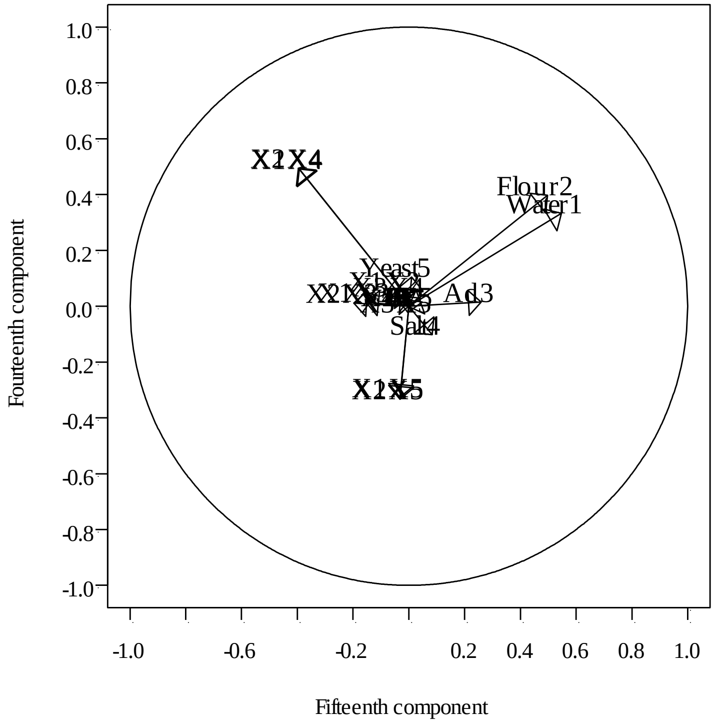
Singular values of W	2.20	0.34	0.17	0.029	0.017

Factorisation 2: collinearity constraints



Component	Singular values
1	2.84
2	1.95
3	1.43
4	1.02
5	0.19
6	0.04
7	0.03
8	0.02
9	0.018
10	0.01
11	2.8 E-16 ~ 0
12	2.6 E-16 ~ 0
13	2.2 E-16 ~ 0
14	1.5 E-16 ~ 0
15	4.8 E-17 ~ 0

Factorisation 2: collinearity constraints



Component	Singular values
1	2.84
2	1.95
3	1.43
4	1.02
5	0.19
6	0.04
7	0.03
8	0.02
9	0.018
10	0.01
11	2.8 E-16 ~ 0
12	2.6 E-16 ~ 0
13	2.2 E-16 ~ 0
14	1.5 E-16 ~ 0
15	4.8 E-17 ~ 0

Factorisation 2: collinearity constraints

Component	Singular values
11	2.8 E-16 ~ 0
12	2.6 E-16 ~ 0
13	2.2 E-16 ~ 0
14	1.5 E-16 ~ 0
15	4.8 E-17 ~ 0

➡ Corresponding PC are defining the collinearity constraints

	PC11	PC12	PC13	PC14	PC15
Water1	-0.29	-0.06	0.44	0.33	0.54
Flour2	0.09	-0.47	-0.22	0.39	0.49
Ad3	0.55	0.27	0.06	0.01	0.25
Salt4	-0.03	0.01	-0.04	-0.10	0.08
Yeast5	-0.06	0.13	-0.11	0.10	0.01
X1X2	0.33	-0.36	-0.58	0.05	-0.04
X1X3	-0.32	-0.25	-0.12	0.01	-0.15
X1X4	0.18	-0.07	0.21	0.49	-0.39
X1X5	0.19	-0.43	0.35	-0.32	-0.03
X2X3	-0.48	-0.27	-0.07	0.01	-0.19
X2X4	0.17	-0.06	0.23	0.49	-0.39
X2X5	0.19	-0.44	0.37	-0.33	-0.03
X3X4	0.006	-0.007	0.01	0.03	-0.02
X3X5	0.01	-0.03	0.02	-0.02	-0.003
X4X5	0.001	-0.001	0.001	-0.0002	-0.0008

How to use this information for more ‘optimal’ design?

- Instead of optimizing variance properties of parameters and predicted response under OLS estimation
- Minimize correlational structure
  - Scree plot of singular values
  - Optimize factorisation 1
- Minimize collinearity structure
  - Optimize factorisation 2

Guarantee design in the ‘full’ region in the exchange algorithm

- Protected design points, fi vertices
- Adapt constraints ⇔ Hat matrix decomposition for leverage studies

Other solutions for the bread example

- Classical solutions
  - Categorisation of the components
  - Define small ranged components as process variables
  - Model reduction
  - Adapted models

Computational constraints

Flare data: 27 candiate points

Calculate all possible designs of 15 points

$$\binom{27}{15} = 139070880$$

- With 200 design evaluations per sec for a quadratic model, it will take 193 h or 8 days

- Thus excursion,branch and bound adaptations

Conclusion

- Plenty of interesting ideas for better mixture design optimality criteria, but computational impossibility to calculate for realistic problems in industry