Construction of Permutation Mixture Experiment Designs

Ben Torsney and Yousif Jaha
University of Glasgow
bent@stats.gla.ac.uk yousif@stats.gla.ac.uk

1 Background

- Mixture Experiments
- Canonical Polynomials (Sheffe' 1958)
- Permutation Mixture Experiment Designs

2 Standard Regression Designs

- Information Matrix $M(\underline{p})$
- D-optimal criteria $\phi(\underline{p}) = logdet(M(\underline{p}))$

3 Permutation Mixture Experiment Designs

- Information Matrix $N(\underline{p})$
- D-optimal criteria $\phi(\underline{p}) = logdet(N(\underline{p}))$

4 General Optimisation Problem (GOP), Optimality Conditions

- PMED as a special case of GOP
- Directional Derivative and Vertex Directional Derivative
- Optimality Conditions

5 Algorithm

6 Constraints under PMED

6.1 Order constraints

- Multiple Local Maxima
- Transformation

6.2 Bound constraints

6.2.1 Common Lower Bound Only

6.2.2 Common Upper Bound Only

6.2.3 Common Lower and Upper Bound

6.3 Simultaneous Order and Bound Constraints

6.3.1 Common Lower Bounds Only and Order Constraints

6.3.2 Common Upper or (Lower and Upper) Bounds and Order Constraints

7 Example

1 Background :

Mixture Experiments :

Any experiment consists of :

control variables	x_u,
response variables	y_u
errors variables	e_u

Mixture experiments with q-components and n blends $(x_{u1}, ..., x_{uq})$ have two extra conditions,

$$\sum_{i=1}^{q} x_{ui} = 1,\ x_{ui} \geq 0,\ i = 1, 2, \cdots, q\ and\ u = 1, ..., n.$$

Examples :

Ceramic Products, Concrete Products, Chemical Products, Rubber Products, Food Products

Canonical Polynomials (Scheffe's Model):

y_u observed at $\mathbf{x}_u = (x_{u1}, ..., x_{uq})$,
$\sum_{i=1}^{q} x_{ui} = 1 \quad and \quad x_{ui} \geq 0, \quad u = 1, 2, ..., n$

Conventional model is quadratic , a good approximation to most functions. The general 2nd order polynomial is:

$$E(y_u) = \alpha + \sum_{i=1}^{q} \beta_i x_{ui} + \sum_{i=1}^{q} \beta_{ii} x_{ui}^2 + \sum_{i\ <}^{q-1} \sum_{j}^{q} \beta_{ij} x_{ui} x_{uj} \quad (1)$$

By making in (1) the substitution :

$$x_j = 1 - \sum_{i \neq j}^{q} x_i$$

the resulting model is the Canonical polynomial

$$E(y_u) = \sum_{i=1}^{q} \beta_i x_{ui} + \sum_{i\ <}^{q-1} \sum_{j}^{q} \beta_{ij} x_{ui} x_{uj}$$

Permutation Mixture Experiment Design :

$$\begin{aligned} \text{Design Points} &= \text{Permutations} \\ \mathbf{x}_u = (x_{u1}, ..., x_{uq}) &= \text{any design point} \\ (x_{u1}, ..., x_{uq}) &= Perm(p_1, ..., p_q) \end{aligned}$$

where $\sum_{i=1}^{q} p_i = 1$ and $p_i \geq 0$

So

$$\mathbf{x}_u = \mathbf{x}_u(\underline{p}) \quad \text{where} \quad \underline{p} = (p_1, \cdots, p_q)^T$$

2 Standard Regression Designs :

Information Matrix is :

$$\mathbf{M}(\underline{p}) = \sum_{i=1}^{k} p_i \mathbf{f}(\mathbf{x}_i)\mathbf{f}^T(\mathbf{x}_i)$$

Where $\mathbf{f}$ is the regression function vector.
D-optimal criteria:

$$\phi(\underline{p}) = \psi(M(\underline{p})) = logdet(M(\underline{p}))$$

3 Permutation Mixture Experiment Designs :

Information Matrix is :

$$\mathbf{N}(\underline{p}) = \sum_{u=1}^{n} \mathbf{f}(\mathbf{x}_u(\underline{p}))\mathbf{f}^T(\mathbf{x}_u(\underline{p})).$$

Where $\mathbf{f}$ is the regression function vector.

$$\mathbf{f}(\mathbf{x}) = (x_1, .., x_q, x_1x_2, .., x_{q-1}x_q)^T$$

D-optimal criteria:

$$\phi(\underline{p}) = \psi(N(\underline{p})) = logdet(N(\underline{p}))$$

4 GOP, Optimality Conditions, Algorithm :

4.1 GOP : maximise $\phi(\underline{p})$
subject to $\sum_{i=1}^{q} p_i = 1$ and $p_i \geq 0 \ i = 1, \cdots, q$

PMED is a special case of the GOP stated above, we seek to optimise

$$\phi(\underline{p}) = \psi\{N(\underline{p})\} = log\{det[N(\underline{p})]\}$$

with respect to $\underline{p}$ (the component values) subject to $p_i \geq 0$, $\sum_{i=1}^{q} p_i = 1$.

4.2 Definition (Directional Derivative) :

Define $g(\underline{p}, \underline{z}, \epsilon)$ as follows :

$$g(\underline{p}, \underline{z}, \epsilon) = \phi\{(1-\epsilon)\underline{p} + \epsilon\underline{z}\}$$

Then the directional derivative of $\phi(.)$ at $\underline{p}$ in the direction of $\underline{z}$ $F_\phi\{\underline{p}, \underline{z}\}$ is defined as follows:

$$\begin{aligned} F_\phi\{\underline{p}, \underline{z}\} &= \lim_{\epsilon \downarrow 0} \frac{\phi\{(1-\epsilon)\underline{p} + \epsilon\underline{z}\} - \phi(\underline{p})}{\epsilon} \\ &= \lim_{\epsilon \downarrow 0} \frac{g(\underline{p}, \underline{z}, \epsilon) - g(\underline{p}, \underline{0}, 0)}{\epsilon} = \left.\frac{dg(\underline{p}, \underline{z}, \epsilon)}{d\epsilon}\right| \epsilon = 0^+ \end{aligned}$$

Whittle (1973) called F_ϕ the directional derivative of $\phi(.)$ at $\underline{p}$ in the direction of $\underline{z}$. This derivative exists even if $\phi(.)$ is not differentiable. If $\phi(.)$ is differentiable then

$$F_\phi(\underline{p}, \underline{z}) = (\underline{p} - \underline{z})^T \frac{\partial \phi}{\partial \underline{p}} = \sum_{i=1}^{q} (p_i - z_i) d_i \tag{2}$$

where $d_i = \frac{\partial \phi}{\partial p_i} \quad i = 1, 2, \cdots, q$. If we substitute the unit vector $\underline{e}_i$ for $\underline{z}$ then we have a vertex directional derivative F_j i.e

$$F_j = F(\underline{p}, \underline{e}_j) = d_j - \sum_{i=1}^{q} p_i d_i. \tag{3}$$

We employ the directional derivative of $\phi(\underline{p})$ to determine necessary first order optimality conditions.

4.3 Condition for local optimality :

If $\phi(.)$ is differentiable at $\underline{p}^*$, then a necessary condition for $\phi(\underline{p}^*)$ to be a local maximum of $\phi(.)$ in the feasible region of GOP is

$$F_j^* = F_\phi\{\underline{p}^*, \underline{e}_j\} = \begin{cases} = 0 & : \quad if\ \underline{p}_j^* > 0 \\ \leq 0 & : \quad if\ \underline{p}_j^* = 0 \end{cases}$$

This is also a sufficient condition for a global maximum of the GOP if $\phi(.)$ is a concave function on its feasible region, as is the case with standard linear regression design problems.

5 Multiplicative Algorithm :

Maximise $\phi(\underline{p}) = logdet(N(\underline{p}))$,

Subject to $\sum_{i=1}^{q} p_i = 1 \qquad and \qquad p_i \geq 0$

The following iteration was used to solve the above problem

$$p_k^{(r+1)} = \frac{p_k^{(r)} G(F_k^{(r)})}{\sum_{i=1}^{q} p_i^{(r)} G(F_i^{(r)})} = H_k(\underline{p}^{(r)}) \tag{4}$$

Where

$$F_k^{(r)} = \frac{\partial \phi^{(r)}}{\partial p_k} - \sum_{i=1}^{q} p_i^{(r)} \frac{\partial \phi^{(r)}}{\partial p_i} \quad ;$$

$$G(F) > 0 \quad and \quad G'(F) > 0$$

We take $G(.) = \Phi(.)$

6 Constraints Under PMED :

6.1 Order Constraints :

Multiple Local Maxima :
Find maximum subject to given ordering of $p_1, \cdots, p_q$.

$$e.g. \qquad p_1 < p_2 < \cdots < p_q.$$
$$or \qquad p_{i_1} < p_{i_2} < \cdots < p_{i_q}$$

Transformations :

1^{st} : $u_1 = p_{i_1},\ u_j = p_{i_j} - p_{i_{j-1}},\ j = 2, \cdots, q, \qquad u_j \geq 0$

$$1 = \sum_{j=1}^{q} p_{i_j} = \sum_{j}^{q} c_j u_j \quad \text{where} \qquad c_j = q - j + 1$$

2^{nd} :
$$s_j = c_j u_j \qquad s_j \geq 0$$
$$\sum_{j=1}^{q} s_j = 1$$

Algorithm :

$$s_k^{(r+1)} = H_k(s^{(r)}) \qquad \text{(see (4))}$$
$$s_k^{(0)} = \frac{1}{q}$$

6.2 Bound Constraints :

6.2.1 Common Lower Bound Only (l):

Note that $0 \leq l \leq \frac{1}{q}$. The feasible region is a convex polyhedron whose vertices are of the form

$$perm(l, \cdots, l, c_0), \quad \text{where} \quad c_0 = 1 - (q-1)l$$

6.2.2 Common Upper Bound Only (u) :

Case I : If u is in the j^{th} interval

$$\frac{1}{q-(j-1)} < u < \frac{1}{q-j}, \; j = 1, \cdots, q-1,$$

then u satisfies $(q-(j-1)u + c_{q-j} = 1$ where $c_{q-j} = 1-(q-j)u$.
The vertices have the following form

$$perm(\overbrace{u, \cdots, u}^{(q-j)\ times}, \overbrace{0, \cdots, 0}^{(j-1)\ times}, c_{q-j}).$$

Notice that the total number of these points is $n_u = \frac{q!}{(j-1)!(q-j)!}$. In general, the vertices will be on the boundary of the simplex except when $j = 1$.

Case II : If $u = \frac{1}{q-j}, j = 1, \cdots, q-2$, then the vertices have the form

$$(\overbrace{u, \cdots, u}^{(q-j)\ times}, \overbrace{0, \cdots, 0}^{j\ times}).$$

6.2.3 Common Lower and Upper Bound :

For problems with simultaneous common lower and upper bounds i.e. constraints of the type $l \leq p_i \leq u$, where $0 < l < \frac{1}{q} < u < 1$, the feasible region is again a regular convex polyhedron with one of the following types of vertex.

Type (1) : For $i = 0, \cdots, q$

$$perm(\overbrace{u, \cdots, u}^{i\ times}, \overbrace{l, \cdots, l}^{q-i\ times})$$

where $iu + (q-i)l = 1$, (Note this includes the possibility of $i = 0$ and $i = q$)

Type (2) : For $i = 0, \cdots, q-1$

$$perm(\overbrace{u, \cdots, u}^{i\ times}, \quad \overbrace{l, \cdots, l}^{(q-i-1)\ times} \quad, c_i),$$

where $c_i = 1 - \{(q-i-1)l + iu\}$ and $(q-i-1)l + iu < 1$

If (u, l) lies on the line $iu + (q-i)l = 1$ for $i = 1, \cdots, q$ then the vertices are of the first type. Note that in the cases $i = 0$ and $i = q$ there is only one vertex namely the common blend $(\frac{1}{q}, \frac{1}{q}, \cdots, \frac{1}{q})$ (viewed as $l = \frac{1}{q}$ when $i = 0$ and $u = \frac{1}{q}$ when $i = q$). This is the only feasible mixture.
Otherwise (u, l) must for some i ($i = 0, \cdots, q-1$) lie on or between the lines $iu + (q-i)l = 1$ and $(i+1)u + (q-i-1)l = 1$.

Note that in the cases $i = 0$ and $i = q - 1$ the vertices are respectively of the forms

$$perm(l, \cdots, l, c_0), \text{ where } c_0 = 1 - (q-1)l$$

and

$$perm(u, \cdots, u, c_{q-1}), \text{where } c_{q-1} = 1 - (q-1)u.$$

Since $c_0 < u$ and $c_{q-1} > l$ under the relevant conditions c_0 and c_{q-1} are in effect revised upper and lower bounds respectively in these two extreme cases.

6.3 Simultaneous Order and Bound Constraints :

6.3.1 Common Lower Bound and Order Constraints :

This is a variation on case 6.1 but with $u_1 = p_{i_1} - l$. So we linearly transform the variables $p_1, \cdots, p_q$ to variables $s_1, \cdots, s_q$ leading to another example of GOP. i.e.

$$\underline{s} = B\underline{p}, \tag{5}$$

where B is a non singular square matrix of order q, AND

$$\sum_{i=1}^{q} s_i = 1 \text{ and } s_i \geq 0;\ i = 1, \cdots, q \tag{6}$$

are automatically satisfied. Hence, it is an example of GOP. Can be solved using the Multiplicative algorithm.

6.3.2 Common Upper or (Upper and Lower) Bound and Order Constraints :

Similarly, we transform the variables $p_1, \cdots, p_q$ to $t_1, \cdots, t_{q+1}$, which must be nonnegative and satisfy 2 equations :

$$\underline{1}^T H(u-l)\underline{t} = \underline{1}^T H \underline{b}$$

$$\underline{1}^T \underline{t} = 1$$

where $H = (B^T B)^{-1} B^T$, $\underline{b} = (-l, 0, \cdots, 0, u)^T$ and

$$B = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & -1 \end{pmatrix} \quad (7)$$

So $\underline{t}$ must lie in a convex polyhedron whose vertices are the BFS of the above system of equations i.e.

$$\underline{t} = \sum_{i=1}^{m} \lambda_i \underline{v}_i \quad (8)$$

for some $\lambda_1, \cdots, \lambda_m$, substituting $\lambda_i \geq 0$, $\sum_i^m \lambda_i = 1$.

So we have transformed to a problem of optimising the D-criterion with respect to the convex weights $\lambda_1, \cdots, \lambda_m$ of the vertices $\underline{v}_1, \cdots, \underline{v}_m$. Again this is a special case of GOP which can be solved using the Multiplicative algorithm.

7 Example :

Assume a design for a 7-mixture experiment consists of 66 design points 63 are subsets of all possible permutations of the fixed proportions $(p_1, p_2, \cdots, p_7)$ and 3 are replicates of the point $(\frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7})$. Assume the following constraints:

(1) Bound constraints : $l = 0.1, u = 0.4$, where l is the common lower bound and u the common upper bound. Note that with just these constraints the feasible region has vertices

$$Perm(0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.4)$$

(2) Order Constraints : in addition suppose we want to find the D-optimal value subject to $p_6 \leq p_3 \leq p_1 \leq p_7 \leq p_2 \leq p_5 \leq p_4$. The vertices of the feasible region in terms of the transformed variables $\underline{t} = (t_1, t_2, \cdots, t_7)$ are :

$\underline{\tilde{v}}_1 = (0,0,0,0,0,0,1,0)$, $\underline{\tilde{v}}_2 = (0,0,0,0,0,0.5,0,.5)$,
$\underline{\tilde{v}}_3 = (0,0,0,0,\frac{1}{3},0,0,\frac{2}{3})$, $\underline{\tilde{v}}_4 = (0,0,0,\frac{1}{4},0,0,0,\frac{3}{4})$,
$\underline{\tilde{v}}_5 = (0,0,0.2,0,0,0,0,.8)$, $\underline{\tilde{v}}_6 = (0,\frac{5}{6},0,0,0,0,0,\frac{1}{6})$,
$\underline{\tilde{v}}_7 = (\frac{1}{7},0,0,0,0,0,0,\frac{6}{7})$.

In terms of p's they are:

$\underline{v}_1 = (.1, .1, .1, .4, .1, .1, .1)$,
$\underline{v}_2 = (.1, .1, .1, .25, .25, .1, .1)$,
$\underline{v}_3 = (.1, .2, .1, .2, .2, .1, .1)$,
$\underline{v}_4 = (.1, .175, .1, .175, .175, .1, .175)$,
$\underline{v}_5 = (.16, .16, .1, .16, .16, .1, .16)$,
$\underline{v}_6 = (.35, .35, .35, .35, .35, .1, .35)$,
$\underline{v}_7 = (\frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7},)$.

The D-optimal solution is $p_1 = p_2 = p_3 = 0.1$, $p_4 = 0.304485$, $p_5 = 0.195515$, $p_6 = p_7 = 0.1$. These were found using the multiplicative algorithm. The following Table shows the local D-optimal values and D-efficiencies subject to the same common lower and upper bounds $l = .1, u = .4$ and to 12 different order constraints. With 2 or 3 exceptions D-efficiencies are all relatively high.

No.	$(p_1, p_2, p_3, p_4, p_5, p_6, p_7)$	$\sqrt[28]{D^*}$	D-efficiency
1	(.1,.1,.1,.30541,.19459,.1,.1)	.0025	75.73%
2	(.1,.30541,.1,.1,.1,.19459,.1)	.0032	100 %
3	(.1,.19459,.1,.1,.1,.30541,.1)	.0032	100 %
4	(.1,.19459,.1,.1,.30541,.1,.1)	.0029	90.96 %
5	(.1,.1,.1,.1,.1,.19459,.30541)	.0029	90.96 %
6	(.1,.1,.1,.19459,.1,.1,.30541)	.0029	90.96 %
7	(.30541,.1,.1,.1,.1,.19459,.1)	.0029	90.96%
8	(.1,.1,.19459,.1,.30541,.1,.1)	.0029	90.96 %
9	(.19459,.1,.30541,.1,.1,.1,.1)	.0032	100 %
10	(.1,.19459,.1,.1,.1,.30541,.1)	.0032	100 %
11	(.290937,.154532,.154532,.1,.1,.1,.1)	.0018	55.50 %
12	(.1,.1,.1,.283948,.129722,.186329,.1)	.0015	45.64 %