**\section{Information decomposition of designed experiments}**

**\label{sec:infoDecomp}**

An experiment's design is intimately linked to the linear model which describes it. More specifically, the linear model describes the relationships between the experimental units (i.e. \emph{block structure}), the relationships between the treatments, (i.e. \emph{treatment structure}), and the assignment of treatments to the experimental units. The following considers an arbitrary design involving both block and treatment factors whose relationships can be represented mathematically by a linear \emph{mixed-effects model}. We show how the raw data from such an experiment are decomposed into their constituent components based on their block and treatment structures.

**\subsection{The linear mixed-effects model}**

**\label{subsec:matrixLMM}**

Consider an experiment involving $n$ observations and $v$ treatment factors, where $t\_i$ denotes the number of treatments in the $i$th treatment factor, and $b$ block factors with $m\_j$ denote the number of blocks in the $j$th block factor, $i = 1,2, \dots, v; j = 1,2,\dots, b$. The linear mixed-effects model for such an experiment can be written in matrix notation as

\begin{equation}\**label{eq:matrixLMM}**

\bm{y} = \mathbf{1}\mu + \X \bm{\alpha} + \Z\bm{\beta} + \bm{\epsilon},

\end{equation}

where $\bm{y}$ is an $n \times 1$ vector of responses, $\mathbf{1}$ is an $n \times 1$ vector whose elements are all unity, $\mu$ denotes the grand mean of the data, and $\bm{\epsilon}\sim \mathcal{N}(0,\sigma^2 \I\_n)$ is a $n \times 1$ vector of unobserved random experimental errors, where $\I\_n$, or $\I$ for clarity, denotes the $n \times n$ identity matrix. The treatment parameter vector is thus defined as

\begin{equation}

\**label{eq:treatPar}**

\bm{\alpha} = (\alpha\_{11 \dots 1}, \alpha\_{11 \dots 2}, \dots, \alpha\_{1 t\_2 \dots t\_t},\dots,\alpha\_{t\_1 t\_2 \dots t\_v}),

\end{equation}

where $\alpha\_{x\_1 x\_2 \dots x\_v}$ denotes the effects from treatment combination $f\_1 f\_2 \dots f\_v$, where $f\_i = 1, \dots, t\_i; i = 1,2, \dots, v$. The treatment design matrix, $\X$, in (\ref{eq:matrixLMM}) defines the allocation of treatment combinations to the experimental and/or observational units. The observational unit is the smallest unit of the experiment. The dimension of $\X$ comprises $n$ rows, and corresponds to the number of observations, while $\prod^v\_{i = 1} t\_i$ comprises columns, and corresponds to the parameter in $\bm{\alpha}$. The vector of block parameters is defined as

\begin{equation}\**label{eq:block1Par}**

\bm{\beta} = (\bm{\beta}\_1, \bm{\beta}\_2, \ldots, \bm{\beta}\_b),

\end{equation}

where

\[

\bm{\beta}\_j = (\beta\_{j1}, \beta\_{j2}, \dots, \beta\_{j m\_j})

\]

and

$\beta\_{jk} ~ \sim \mathcal{N}(0, \sigma\_j^2)$ ($j=1,2,\dots,b; k=1,2,\dots, m\_j$). The block design matrix, $\Z$, in (\ref{eq:matrixLMM}) can then be expressed as

\begin{equation}\**label{eq:block1Mat}**

\Z = [\Z\_1 \vert \Z\_2 \vert \ldots \vert \Z\_b],

\end{equation}

where $\Z\_j$ is the design matrix corresponding to $\bm{\beta}\_j$. Thus, the dimension of $\Z$ consists of $n$ rows, and corresponds again to the number of observation, and the $\sum^{b}\_{j = 1} m\_j$ columns, and corresponds to the parameter in $\bm{\beta}$. The treatment and block design matrices presented in this section are shown to be essential components of the decomposition method described in the remainder of this section.

**\subsection{Null decomposition using projection matrices}**

**\label{subsec:strataDecompProj}**

The vector of response, $\bm{y}$, in (\ref{eq:matrixLMM}) spans an $n$-dimensional Euclidean space, commonly denoted by $\mathbb{R}^n$. A vector space, $\mathbb{V}$, is a \emph{subspace} of $\mathbb{R}^n$, i.e.\ $\mathbb{V} \subset \mathbb{R}^n$, if every vector in $\mathbb{V}$ is also in $\mathbb{R}^n$ \citep{Hadi1996}. The information decomposition of $\bm{y}$ involves its separation from the $\mathbb{R}^n$ space into its constituent vector subspace components. These vector subspaces correspond to what are commonly referred to as the \emph{strata} of the ANOVA, and can be mathematically expressed as

\begin{equation}

\**label{eq:vecSpace}**

\mathbb{R}^n = \mathbb{V}\_0 \oplus \mathbb{V}\_1 \oplus, \dots , \oplus \mathbb{V}\_{b} \oplus \mathbb{V}\_{b + 1},

\end{equation}

where $\oplus$ denotes the addition operator of the vector spaces, and $\mathbb{V}\_j$ denotes the $j$th stratum that corresponds to the $j$th block parameters and block design sub-matrices defined in (\ref{eq:block1Par}) and (\ref{eq:block1Mat}), respectively. For any designed experiment, the first ad last elements, $\mathbb{V}\_0$ and $\mathbb{V}\_{b + 1}$, in (\ref{eq:vecSpace}) always denote the vector subspace for the grand mean and experimental error, respectively.

Since the decomposition is the separation of the known variation in the data, the variance structure of the data, $\bm{y}$, can be expressed in a spectral form as

\begin{equation}

\**label{eq:strata}**

\operatorname{Var}(\bm{y}) = \sum\_{j=0}^{b + 1} \xi\_j \Q\_j,

\end{equation}

where $\Q\_j$ is an $n \times n$ matrix, i.e.\ $\Q\_j^2 = \Q\_j$, and thus is the \emph{orthogonal projector} of $\bm{y}$ onto the vector subspace $\mathbb{V}\_j$ (i.e.\ stratum $j$) and $\xi\_j$ is the $j$th stratum variance (i.e.\ $\xi\_j = \operatorname{Var}(\Q\_j \bm{y})$). Since the orthogonal projectors are used to project a vector from one vector subspace to another, orthogonal projectors are \emph{projection matrices} which are symmetric, i.e.\ $\Q\_j' = \Q\_j$, orthogonal, i.e.\ $\Q\_a\Q\_b = 0$ and idempotent, i.e.\ $\Q\_j^2 = \Q\_j$ \citep{Hadi1996}.

Since $\mathbb{V}\_0$ and $\mathbb{V}\_{b + 1}$ are the vector subspaces for the grand mean and the experimental error, respectively, $\Q\_0 \bm{y})$ and $\Q\_{b+1} \bm{y})$ are the estimates for the grand mean, $\mathbf{1}\mu$, and experimental error, $\bm{\epsilon}$, respectively, in (\ref{eq:matrixLMM}). Furthermore, the vector $\Z\_j\bm{\beta}\_j$, which contains the $j$th vector of the block design matrix and parameter, can be estimated from $\Q\_j \bm{y})$. Thus, the SS of $\bm{y}$, i.e.\ $\bm{y}'\bm{y}$, can be decomposed into $b + 2$ components of the SS, i.e.\

\begin{equation}

\**label{eq:decomp}**

\bm{y}'\bm{y} = \sum\_{j=0}^{b + 1}\bm{y}'\Q\_j\bm{y},

\end{equation}

where $\bm{y}'\Q\_j\bm{y}$ denotes the total SS in the $j$th stratum. Equations~(\ref{eq:vecSpace}), (\ref{eq:strata}) and (\ref{eq:decomp}) thus give a basic illustration of the decomposition of the data vector ignoring the treatment, which we refer to as the \emph{null decomposition}. The process results from the sequential fitting of the block factors, and from the computation of the SS of each block factor and the subtraction of the computed SS from the total SS, a process also known as \emph{sweeping} introduced by \cite{Wilkinson1970} and \cite{Payne1977} \citep{Brien1999}. The remainder of this section describes each step of the null decomposition, particularly in terms of how the orthogonal projector, denoted by $\Q\_j$, is computed in stratum $j$.

For any designed experiment, the initial step of null decomposition is to sweep the grand mean from the data vector, $\bm{y}$. Since $\mu$, in (\ref{eq:matrixLMM}), is a vector of length $1$, the grand mean vector spans a $1$-dimensional grand mean vector subspace, denoted by $\mathbb{V}\_0$ in (\ref{eq:vecSpace}). To sweep the grand mean from $\bm{y}$, $\bm{y}$ is first projected onto $\mathbb{V}\_0$, which is pre-multiplied by $\Q\_{0}$, giving $\Q\_{0}\bm{y}$. The $\Q\_{0}$ is given by

\begin{equation}

\**label{eq:vectorProj}**

\Q\_{0} = \mP\_{\K} = {\K}({\K}'{\K})^{-1}{\K}',

\end{equation}

where $\mP\_K$ denotes the projection matrix of matrix $\K$, which is an $n \times n$ averaging matrix with all elements equal to ${n}^{-1}$. Since $\mP\_{\K}$ can be shown be the same as $\K$, for clarity for the rest of this section, $\mP\_{\K} \bm{y}$ can be rewritten as $\K \bm{y}$.

The next step is to obtain the orthogonal complement of $\K\bm{y}$ by subtracting the $\K\bm{y}$ from $\bm{y}$, i.e.\

\[\bm{y} - \K\bm{y} = (\I-\K)\bm{y},\]

where $(\I-\K)\bm{y}$ denotes the \emph{mean corrected observational vector}, which spans $\mathbb{V}^{\perp}\_0$ with dimension of $(n - 1)$. The vector subspace $\mathbb{V}^{\perp}\_0$ is the \emph{orthogonal complement} of $\mathbb{V}\_0$. Furthermore, the adjusted total SS is obtained by pre-multiplying the $(\I-\K)\bm{y}$ by its transpose, i.e.\

\begin{equation}

\**label{eq:adjustSS}**

[(\I-\K)\bm{y}]'[(\I-\K)\bm{y}] = \bm{y}'(\I-\K)\bm{y}.

\end{equation}

The total adjusted SS can be calculated by subtraction

\[

\bm{y}'\bm{y} - \bm{y}'\K\bm{y} = \bm{y}'(\I-\K)\bm{y}.

\]

The vector $(\I-\K)\bm{y}$ is then projected onto the vector subspace of the stratum $1$, i.e.\ $\mathbb{V}\_1$, which results in $\Q\_{1}\bm{y}$, where $\Q\_{1}$ is the orthogonal projector of the stratum that corresponds to the first block parameter, $\bm{\beta}\_1$. Providing the $\mP\_{\Z\_1}$ is the projection matrix of $\Z\_1$, the orthogonal projector $\Q\_{1}$ is given by the $\I-\K$ pre-multiplied by $\mP\_{\Z\_1}$. Thus, the vector $\Q\_{1}\bm{y}$ can be re-written as

\begin{equation}\**label{eq:projectB}**

\mP\_{\Z\_1}[(\I-\K)\bm{y}] = (\mP\_{\Z\_1} - \K)\bm{y},

\end{equation}

where vector $(\mP\_{\Z\_1} - \K)\bm{y}$ represents estimates of block effects in the vector $\bm{\beta}\_1$. The orthogonal complement of $(\mP\_{\Z\_1} - \K)\bm{y}$ is derived by subtraction from the vector $(\I-\K)\bm{y}$ as

\begin{equation}

\**label{eq:orthComp}**

(\I-\K)\bm{y}- (\mP\_{\Z\_1} - \K)\bm{y} = (\I -\mP\_{\Z\_1})\bm{y},

\end{equation}

which corresponds to the elimination of the effects in $\bm{\beta}\_1$.

The SS are derived by pre-multiplying the vectors in (\ref{eq:projectB}) and (\ref{eq:orthComp}) by their transpose, as described in (\ref{eq:adjustSS}), i.e.\

\[

\bm{y}'(\I-\K)\bm{y}- \bm{y}'(\mP\_{\Z\_1} - \K)\bm{y} = \bm{y}'(\I -\mP\_{\Z\_1})\bm{y}.

\]

If the block structure contains additional block factors, e.g.\ plots and/or subplots, the vector $(\I - \mP\_{\Z\_1})\bm{y}$ is further projected onto the next vector subspace, $\mathbb{V}\_2$. Thus, in general, the projection of the data vector, $\bm{y}$ from $\mathbb{V}\_{j}$ onto $\mathbb{V}\_{j + 1}$ can be written as $\mP\_{\Z\_{j+1}}\Q\_{j}\bm{y}$. The orthogonal complement of $\mP\_{\Z\_{j+1}}\Q\_{j}\bm{y}$ can be derived by subtraction, i.e.\

\begin{equation}

\**label{eq:orthCompSummary}**

\Q\_{j}\bm{y}- \mP\_{\Z\_{j+1}}\Q\_{j}\bm{y} = (\I -\mP\_{\Z\_{j+1}})\Q\_{j}\bm{y} = \Q\_{j+ 1}\bm{y}, \; (j = 0,1,\dots, b, b+1; \Q\_0 = \K).

\end{equation}

The SS thus are derived by pre-multiplying the vectors in (\ref{eq:orthCompSummary}) as

\[

\bm{y}'\Q\_{j}\bm{y}- \bm{y}'\Q\_{j}\mP\_{\Z\_{j+1}}\Q\_{j}\bm{y} = \bm{y}'\Q\_{j}(\I -\mP\_{\Z\_{j+1}})\Q\_{j}\bm{y}=\bm{y}'\Q\_{j+ 1}\bm{y}, \; (j = 0,1,\dots, b, b+1; \Q\_0 = \K).

\]

Provided the SS for each stratum is defined using the orthogonal projectors, the EMS can be computed for the theoretical ANOVA table. From (\ref{eq:decomp}), the expected sum of squares (ESS) of the $i$th stratum without the treatment effects can be shown as

\begin{equation}

\**label{eq:ESSQuad}**

\operatorname{E}(\bm{y}'\Q\_i\bm{y})= \mathrm{tr}(\Q\_i)\operatorname{cov}(\bm{y}),

\end{equation}

where $\operatorname{cov}(\bm{y})$ is the variance covariance matrix and $\mathrm{tr}(\Q\_i)$ is the \emph{trace} of the matrix $\Q\_i$ \citep{Searle1982}.

Consider an experiment arranged in RCBD, then the null decomposition of the total SS can be expressed as

\begin{equation}

\**label{eq:infoDecomp1}**

\bm{y}'\bm{y} = \bm{y}'\K\bm{y} + \bm{y}'(\mP\_{\Z\_1}-\K)\bm{y} + \bm{y}'(\I - \mP\_{\Z\_1})\bm{y},

\end{equation}

where $\K$, $(\mP\_{\Z\_1}-\K)$ and $(\I - \mP\_{\Z\_1})$ denote the orthogonal projectors of the grand mean, Between and Within Blocks strata, i.e.\ $\mathbb{V}\_0$, $\mathbb{V}\_1$ and $\mathbb{V}\_2$, respectively.

Since the SS of the Between Blocks stratum is $\bm{y}'(\mP\_{\Z\_1}-\K)\bm{y}$, the ESS of the Between Blocks stratum can be shown as

\begin{eqnarray}

\nonumber \operatorname{E}(\bm{y}'(\mP\_{\Z\_1}-\K)\bm{y}) &=& \mathrm{tr}(\mP\_{\Z\_1}-\K)\operatorname{cov}(\bm{y})\\

\nonumber &=& \mathrm{tr}(\mP\_{\Z\_1}-\K)\operatorname{cov}(\bm{\epsilon}) + \mathrm{tr}[\Z'(\mP\_{\Z\_1}-\K)\Z]\operatorname{cov}(\bm{\beta})\\

\nonumber &=& (m\_1 - 1)\sigma^2 + (m\_1 - 1)v\sigma^2\_1\\

**\label{eq:computRBD}** &=& (m\_1 - 1)(\sigma^2 +m\_2\sigma^2\_1),

\end{eqnarray}

where $m\_1$ and $m\_2$ denote the block number and block size, respectively. Subsequently, the EMS is calculated by dividing the ESS by the corresponding DF. Hence, the EMS of the Between Blocks stratum is $\sigma^2 +v\sigma^2\_1$. Similarly, the EMS of the Within Blocks stratum can be shown as $\sigma^2$. Table~\ref{tab:infoDecomp} shows the theoretical ANOVA table without the treatment components for RCBD.