MT4614 Design of Experiments

R. A. Bailey
University of St Andrews

26 January 2022

Section 2.2: Treatment subspace (an example)

plot	tmt	typical vector $\mathbf{v}$	data vector $\mathbf{y}$	some vectors in V_T			an obvious basis $\mathbf{u}_A$	$\mathbf{u}_B$	$\mathbf{u}_C$	$\boldsymbol{\tau}$	$\hat{\boldsymbol{\tau}}$
1	A	v_1	y_1	1	1	1/5	1	0	0	τ_A	$\hat{\tau}_A$
2	C	v_2	y_2	3	0	−1/4	0	0	1	τ_C	$\hat{\tau}_C$
3	A	v_3	y_3	1	1	1/5	1	0	0	τ_A	$\hat{\tau}_A$
4	B	v_4	y_4	0	2	0	0	1	0	τ_B	$\hat{\tau}_B$
5	C	v_5	y_5	3	0	−1/4	0	0	1	τ_C	$\hat{\tau}_C$
6	B	v_6	y_6	0	2	0	0	1	0	τ_B	$\hat{\tau}_B$
7	A	v_7	y_7	1	1	1/5	1	0	0	τ_A	$\hat{\tau}_A$
8	A	v_8	y_8	1	1	1/5	1	0	0	τ_A	$\hat{\tau}_A$
9	B	v_9	y_9	0	2	0	0	1	0	τ_B	$\hat{\tau}_B$
10	A	v_{10}	y_{10}	1	1	1/5	1	0	0	τ_A	$\hat{\tau}_A$
11	C	v_{11}	y_{11}	3	0	−1/4	0	0	1	τ_C	$\hat{\tau}_C$
12	B	v_{12}	y_{12}	0	2	0	0	1	0	τ_B	$\hat{\tau}_B$
13	C	v_{13}	y_{13}	3	0	−1/4	0	0	1	τ_C	$\hat{\tau}_C$

Section 2.2: Some vectors in the example

The first treatment vector is $\mathbf{u}_A + 3\mathbf{u}_C$.

The second treatment vector is $\mathbf{u}_A + 2\mathbf{u}_B$.

The third treatment vector is $\frac{1}{5}\mathbf{u}_A - \frac{1}{4}\mathbf{u}_C$.
Its coefficients sum to zero, so it is a treatment contrast.

The vector $\boldsymbol{\tau}$ is the vector of unknown parameters that we wish to estimate.
In fact

$$\boldsymbol{\tau} = \tau_A\mathbf{u}_A + \tau_B\mathbf{u}_B + \tau_C\mathbf{u}_C.$$

After we have the data, we obtain estimates $\hat{\tau}_A$, $\hat{\tau}_B$ and $\hat{\tau}_C$ for these parameters.
These give us the vector of fitted values

$$\hat{\boldsymbol{\tau}} = \hat{\tau}_A\mathbf{u}_A + \hat{\tau}_B\mathbf{u}_B + \hat{\tau}_C\mathbf{u}_C.$$

Break for non-technical stuff

All information about MT4614, including timetable, summary of material covered, problem sheets (when available), data files (when needed), is on the web page
`http://www-groups.mcs.st-and.ac.uk/~rab/MT4614/`
There is a direct link to this from MMS.

If you have not already got it, please download Problem Sheet 1 from there and think about it before this week's tutorial.

I encourage you all to write out the notes of each lecture in your own handwriting, with your own comments and explanations added. This is really the only way that the material will get into your brain.

If you have any questions, do not hesitate to email me at
`rab24@st-andrews.ac.uk`
(but do not expect replies during weekends).

Any questions on this lecture so far?

Section 2.3: Orthogonality

The usual scalar product on V is

$$\mathbf{v} \cdot \mathbf{w} = \sum_{\omega \in \Omega} v_\omega w_\omega = \mathbf{v}'\mathbf{w}.$$

Put

$$\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v} = \text{squared length of } \mathbf{v}.$$

We say that $\mathbf{v}$ is orthogonal to $\mathbf{w}$ if $\mathbf{v} \cdot \mathbf{w} = 0$.
Write this as $\mathbf{v} \perp \mathbf{w}$.

Let W be any subspace of V.
The orthogonal complement of W is

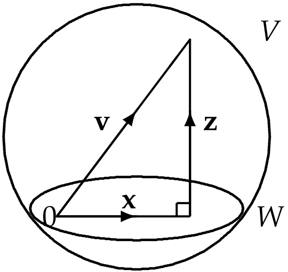
$$W^\perp = \{\mathbf{v} \in V : \mathbf{v} \perp \mathbf{w} \quad \forall \mathbf{w} \in W\}.$$

- $W^\perp$ is also a subspace of V.
- $\dim(W^\perp) = \dim(V) - \dim(W)$.
- $\left(W^\perp\right)^\perp = W$.

(The last two are true when $\dim(V)$ is finite.)

Section 2.3: Orthogonal projection

$V = W \oplus W^\perp$, which is the direct sum of W and $W^\perp$.



This means that, given $\mathbf{v}$ in V,
there is a unique $\mathbf{x}$ in W and $\mathbf{z}$ in $W^\perp$ such that $\mathbf{v} = \mathbf{x} + \mathbf{z}$.
We call $\mathbf{x}$ the orthogonal projection of $\mathbf{v}$ onto W; write $\mathbf{x} = \mathbf{P}_W\mathbf{v}$.
Similarly, $\mathbf{P}_{W^\perp}\mathbf{v} = \mathbf{z} = \mathbf{v} - \mathbf{P}_W\mathbf{v} = \mathbf{v} - \mathbf{x}$.
Pythagoras' Theorem shows that, for fixed $\mathbf{v}$, as $\mathbf{w}$ varies over W then $\|\mathbf{v} - \mathbf{w}\|^2$ is minimized when $\mathbf{w} = \mathbf{P}_W\mathbf{v}$.

Section 2.3: Orthogonal basis

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an orthogonal basis for W then

$$\mathbf{P}_W \mathbf{v} = \sum_{i=1}^{n} \left(\frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \right) \mathbf{u}_i.$$

Therefore $\mathbf{P}_W$ is a linear operator,
so it can be regarded as a matrix of size $\dim(V) \times \dim(V)$.

- The matrix $\mathbf{P}_W$ is symmetric, so that $\mathbf{P}_W' = \mathbf{P}_W$.
- The matrix $\mathbf{P}_W$ is idempotent, so that $\mathbf{P}_W^2 = \mathbf{P}_W$.
- $\operatorname{rank}(\mathbf{P}_W) = \operatorname{trace}(\mathbf{P}_W) = \dim(W)$.

Suppose that $W = V_T$.
A convenient orthogonal basis for V_T is $\{\mathbf{u}_1, \ldots, \mathbf{u}_t\}$, where

$$\mathbf{u}_i(\omega) = \begin{cases} 1 & \text{if } T(\omega) = i \\ 0 & \text{otherwise} \end{cases}$$

We already used this basis in the example at the beginning of this lecture!

Background material for the next section

The next section uses some properties of variance and covariance of random vectors.

In case you have forgotten this material,
you can find some notes about it on the web page,
under the heading **Basics of Random Variables**.

2.4 Linear Model

Putting the random variables Y_ω into a column vector $\mathbf{Y}$ gives

$$\mathbf{Y} = \boldsymbol{\tau} + \mathbf{Z} \qquad \text{where } \boldsymbol{\tau} = \sum_{i=1}^{t} \tau_i \mathbf{u}_i.$$

The simplest assumption on $\mathbf{Z}$ is that

$$\mathbb{E}(\mathbf{Z}) = \mathbf{0} \quad \text{and} \quad \mathrm{Cov}(\mathbf{Z}) = \sigma^2 \mathbf{I}.$$

So

$$\mathbb{E}(\mathbf{Y}) = \boldsymbol{\tau} \in V_T$$

(unknown $\boldsymbol{\tau}$ in known V_T, and we want to estimate $\boldsymbol{\tau}$)
and

$$\mathrm{Cov}(\mathbf{Y}) = \sigma^2 \mathbf{I}$$

(usually σ^2 is unknown).

2.4 Linear Model (continued)

Let W be a subspace of V.
Expectation commutes with all linear operators, so

$$\mathbb{E}(\mathbf{P}_W \mathbf{Y}) = \mathbf{P}_W \left(\mathbb{E}(\mathbf{Y})\right) = \mathbf{P}_W \boldsymbol{\tau}.$$

If $\mathbf{X}$ is any random vector then

$$\mathbb{E}\left(\|\mathbf{X}\|^2\right) = \mathbb{E}\left(\sum_\omega X_\omega^2\right) = \sum_\omega \mathbb{E}\left(X_\omega^2\right)$$

$$= \sum_\omega \left(\mathrm{Var}(X_\omega) + \left(\mathbb{E}(X_\omega)\right)^2\right)$$

$$= \mathrm{trace}\left(\mathrm{Cov}(\mathbf{X})\right) + \|\mathbb{E}(\mathbf{X})\|^2.$$

2.4 Linear Model (continued some more)

$$\boxed{\mathbb{E}\left(\|\mathbf{X}\|^2\right) = \text{trace}\left(\text{Cov}(\mathbf{X})\right) + \|\mathbb{E}(\mathbf{X})\|^2.}$$

Putting $\mathbf{X} = \mathbf{P}_W\mathbf{Y}$, we have $\mathbb{E}(\mathbf{X}) = \mathbf{P}_W\boldsymbol{\tau}$ and

$$\text{Cov}(\mathbf{X}) = \text{Cov}\left(\mathbf{P}_W\mathbf{Y}\right) = \mathbf{P}_W \,\text{Cov}(\mathbf{Y})\mathbf{P}'_W = \mathbf{P}_W\left(\sigma^2\mathbf{I}\right)\mathbf{P}_W = \sigma^2\mathbf{P}_W.$$

So

$$\begin{aligned} \mathbb{E}\left(\|\mathbf{P}_W\mathbf{Y}\|^2\right) &= \text{trace}\left(\sigma^2\mathbf{P}_W\right) + \|\mathbf{P}_W\boldsymbol{\tau}\|^2 \\ &= \sigma^2\,\text{trace}\left(\mathbf{P}_W\right) + \|\mathbf{P}_W\boldsymbol{\tau}\|^2 \\ &= \sigma^2\dim(W) + \|\mathbf{P}_W\boldsymbol{\tau}\|^2. \end{aligned}$$