## Notes on Profile Analysis

Profile analysis provides a relatively simple form of the analysis of repeated measures. Recall that repeated measures analysis deals with multivariate observations $\mathbf{x}=\left[\mathrm{x}_{1}, \mathrm{x}_{2}, \ldots, \mathrm{x}_{\mathrm{p}}\right]^{\prime}$ where each variate is a measurement or determination on essentially the same quantity under $p$ different conditions or treatments. In particular, the $x_{i}$ 's might be measurements at time $t_{i}, i=1, \ldots, p$.
Variable $x_{i}$ is the response under level i of a within-subject factor. As in other situation where there are multiple treatments, you are primarily interested in the differences among the levels of the within-subject factor.
One way to display the data $x_{1}, x_{2}, \ldots, x_{p}$ for a given individual or case is as a plot of $x_{1}, x_{2}, \ldots$, $x_{p}$ against the variable number $1,2, \ldots, p$. This is sometimes called the profile for that case. Similarly you can make a profile plot of the mean vector $\boldsymbol{\mu}$ as a plot of $\mu_{1}, \mu_{2}, \ldots, \mu_{p}$ against $1,2, \ldots, p$.
Here are some examples in which profile analysis might be appropriate.
(i) $x_{1}, x_{2}, \ldots, x_{6}$ are measurements of a person's heart rate $x$ made at 6 fixed times over the course of 24 hours.
(ii) $x_{1}, \ldots, x_{8}$ are the total amounts of milk obtained from a cow in the $1^{\text {st }}, 2^{\text {nd }}, \ldots, 8^{\text {th }}$ week of lactation.
(iii) $x_{1}, x_{2}, x_{3}, x_{4}$ are ratings of 4 brands of root beer made by a taste test panel member.

## Single sample test of within-subject equality of means

In the simplest case, there is a single sample $\mathbf{x}_{1}, \mathbf{x}_{2}, \ldots, \mathbf{x}_{\mathrm{n}}$ from a multivariate population with mean $\boldsymbol{\mu}=\left[\mu_{1}, \mu_{2}, \ldots, \mu_{p}\right]^{\prime}$ and covariance matrix $\boldsymbol{\Sigma}$. When all components of $\mathbf{x}$ measure essentially the same quantity under different conditions, you should view this as a repeated measures factorial design with one within-subject factor with $p$ levels and no amongsubject factor. There is no among-subject factor since all vectors (cases) come from a single population.
Because different variables represent observations at different times or under different conditions, a hypothesis of immediate interest is usually that there are no differences among variate means, that is

$$
\mathrm{H}_{0}: \mu_{1}=\mu_{2}=\ldots=\mu_{\mathrm{p}}
$$

or equivalently

$$
\mathrm{H}_{0}: \boldsymbol{\mu}=\mu \mathbf{1}_{\mathrm{p}} \text { for some } \mu .
$$

For Example (i) this is the hypothesis that the expected heart rate (population average heart rate) is the same at the 6 times.

For Example (iii), this is the hypothesis that the 4 brands do not differ in their mean ratings.
You can approach many problems in repeated measures analysis by defining suitable
contrasts among the variables. For example, define the following p-1 by p contrast matrix (contrasts specified by rows):

$$
\mathrm{C}_{\mathrm{a}}=\left[\begin{array}{cccccc}
1 & -1 & 0 & 0 & \ldots & 0 \\
1 & 0 & -1 & 0 & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
1 & 0 & 0 & 0 & -1 & 0
\end{array}\right]
$$

Then you can express $H_{0}$ as $H_{0}: \mathbf{C}_{1} \boldsymbol{\mu}=\mathbf{0}$, that is $\mu_{1}-\mu_{2}=0, \mu_{1}-\mu_{3}=0, \ldots, \mu_{1}-\mu_{p}=0$.
Suppose $\mathbf{w}$ is the $\mathrm{p}-1$ dimensional random vector

$$
\mathbf{w}=\mathbf{C}_{\mathrm{a}} \mathbf{x}=\left[\mathrm{x}_{1}-\mathrm{x}_{2}, \mathrm{x}_{1}-\mathrm{x}_{3}, \ldots, \mathrm{x}_{1}-\mathrm{x}_{\mathrm{p}}\right]^{\prime}
$$

Then $\mathbf{w}$ has expectation $\mathrm{E}[\mathbf{w}]=\mathbf{C}_{\mathrm{a}} \boldsymbol{\mu}$ and you can express the null hypothesis as

$$
\mathrm{H}_{0 \mathrm{a}}: \mathrm{E}[\mathbf{w}]=\mathbf{0} .
$$

You can test this hypothesis in any of a number of familiar ways, the most appropriate of which will often be a 1-sample Hotelling's $\mathrm{T}^{2}$ :

$$
\mathrm{T}^{2}=\overline{\mathbf{w}}^{\prime} \hat{\mathrm{V}}[\overline{\mathbf{w}}]^{-1} \overline{\mathbf{w}}=\overline{\mathbf{w}}^{\prime}\left(\frac{1}{\mathrm{n}} \mathbf{S}\right)^{-1} \overline{\mathbf{w}}=\mathrm{n}\left(\mathbf{C}_{\mathrm{a}} \overline{\mathbf{x}}\right)^{\prime}\left(\mathrm{C}_{\mathrm{a}} \mathbf{S}_{\mathrm{x}} \mathrm{C}_{\mathrm{a}}{ }^{\prime}\right)^{-1}\left(\mathbf{C}_{\mathrm{a}} \overline{\mathbf{x}}\right)
$$

where $\overline{\mathbf{x}}$ and $\overline{\mathbf{w}}=\mathbf{C}_{\mathrm{a}} \overline{\mathbf{x}}$ are the sample means of $\mathbf{x}$ and $\mathbf{w}$, and $\mathbf{S}_{\mathrm{x}}$ and $\mathbf{S}_{\mathrm{w}}=\mathbf{C}_{\mathrm{a}} \mathbf{S}_{\mathrm{x}} \mathbf{C}_{\mathrm{a}}{ }^{\prime}$ are their sample unbiased variance matrices on $f_{e}=n-1$ degrees of freedom.
Because the dimension of $\mathbf{w}$ is $\mathrm{q}=\mathrm{p}-1$, the null distribution of $\mathrm{T}^{2}$ (assuming the multivariate normality of $\mathbf{x}$ ) is

$$
\left\{\mathrm{qf}_{\mathrm{e}} /\left(\mathrm{f}_{\mathrm{e}}-\mathrm{q}+1\right)\right\} \times \mathrm{F}_{\mathrm{q}, \mathrm{f}_{\mathrm{e}}-\mathrm{q}+1}=\{(\mathrm{p}-1)(\mathrm{n}-1) /(\mathrm{n}-\mathrm{p}+1)\} \times \mathrm{F}_{\mathrm{p}-1, \mathrm{n}-\mathrm{p}+1} .
$$

In large samples this approaches $\chi_{q}{ }^{2}$.
$C_{a}$ is only one of many different contrast matrices you might use. Let $\mathbf{C}$ be any other full rank contrast matrix, that is a p-1 by p matrix with rank $\mathrm{p}-1$ whose rows sum to zero. Symbolically, $\mathbf{C}$ must satisfy $\mathbf{C 1}_{\mathrm{p}}=\mathbf{0}$. Then you can also express $\mathrm{H}_{0}$ as $\mathrm{H}_{0}: \mathbf{C \mu}=\mathbf{0}$. For example,

$$
\mathrm{C}_{\mathrm{b}}=\left[\begin{array}{cccccc}
1 & -1 & 0 & \ldots & 0 & 0 \\
0 & 1 & -1 & \ldots & 0 & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & 0 & \ldots & 1 & -1
\end{array}\right]
$$

is a contrast matrix which you can use to express the same null hypothesis in the form $H_{0 b}$ : $\mu_{1}=\mu_{2}, \mu_{2}=\mu_{1}, \ldots, \mu_{p-1}=\mu_{\mathrm{p}}$, that is there is no difference between adjacent levels of the within subject treatment or factor.
When the variables represent responses at equally spaced quantitative levels of a factor, perhaps time, you might consider using a matrix $\mathbf{C}$ whose rows are given by orthogonal polynomials weights. For $p=4$, this would be

$$
C_{p}=\left[\begin{array}{cccc}
-3 & -1 & 1 & 3 \\
1 & -1 & -1 & 1 \\
-1 & 3 & -3 & 1
\end{array}\right]
$$

Row j of $\mathrm{C}_{\mathrm{p}}$ is related to a null hypothesis concerning the form of polynomical dependence of the mean on the time.

The use of a different $\mathbf{C} \neq \mathbf{C}_{\mathrm{a}}$ might appear to result in a different $\mathrm{T}^{2}$ statistics, say $\tilde{\mathrm{T}}^{2}=$ $\mathrm{n}(\mathbf{C} \overline{\mathbf{x}})^{\prime}\left(\mathbf{C S}_{\mathrm{x}} \mathbf{C}^{\prime}\right)^{-1}(\mathbf{C} \overline{\mathbf{x}})$. However, there always exists a non-singular $\mathrm{p}-1$ by $\mathrm{p}-1$ matrix $\mathbf{A}$ (for example, $\left.\mathbf{A}=\mathbf{C}_{\mathrm{a}} \mathbf{C}^{\prime}\left(\mathbf{C} C^{\prime}\right)^{-1}\right)$, such that $\mathbf{C}_{a}=\mathbf{A C}$. Substituting in the formula for $T^{2}$, and using the identity $\left(\mathbf{A C S}_{\mathrm{x}} \mathbf{C}^{\prime} \mathbf{A}^{\prime}\right)^{-1}=\left(\mathbf{A}^{\prime}\right)^{-1}\left(\mathbf{C S}_{\mathrm{x}} \mathbf{C}^{\prime}\right)^{-1} \mathbf{A}^{-1}$ it easily follows that $\mathrm{T}^{2}=\tilde{\mathrm{T}}^{2}$. Therefore the value of $\mathrm{T}^{2}$ does not depend on the particular choice of $\mathbf{C}$. You should normally a $\mathbf{C}$ that most closely reflects your interests, such as comparisons with the $\mu_{1}$, comparisons between adjacent means, or the form of a polynomial response function.
There are, of course, other tests of $\mathrm{H}_{0}: \mathbf{C} \boldsymbol{\mu}=\mathbf{0}$, such as Bonferronized single sample t-tests based on $\mathrm{w}_{1}, \ldots, \mathrm{w}_{\mathrm{p}-1}$. Although these tests will depend on the choice of $\mathbf{C}$, they still may be appropriate if $\mathbf{C}$ has been chosen on the basis of research goals. For example, when $\mathrm{x}_{1}$ represents a measurement under a control treatment or an initial value at the start of an experiment, your focus might be on the comparisons of each treatment with the control or initial value, that is comparison of $\mu_{j}$ with $\mu_{1}, j=2, \ldots, p$. In that case, a Bonferronized $t$ analysis based on $\mathbf{w}=\mathbf{C}_{\mathrm{a}} \mathbf{x}$, where $\mathbf{C}_{\mathrm{a}}$ is as above might be most appropriate. Or if you have a special interest in the successive changes between $\mu_{j-1}$ and $\mu_{j}$, a Bonferronized analysis based on $\mathbf{w}=\mathbf{C}_{\mathrm{b}} \mathbf{x}$ might be appropriate as it might reveal when any change took place. If you are interested in the form of a possibly curvilinear dependence on time, an analysis based on $\mathbf{w}$ $=\mathrm{C}_{\mathrm{p}}$ might be appropriate.

## Comparison with Randomized Block Design (RCBD)

The situation just described is deceptively similar to a univariate randomized complete block (RCBD) experiment. The n subjects (individual multivariate observations $\mathbf{x}_{\mathrm{i}}$ ) correspond to n blocks, and the within-subject measurements (components of $\mathbf{x}$ ) correspond to to measurements made on "plots" within a block.
If this approach were applicable, you could do an ordinary univariate two-way (blocks and treatments) Analysis of Variance (ANOVA), followed by univariate analysis of treatment means. In MacAnova, if $x$ is the $n$ by $p$ matrix of responses, you might do the ANOVA by

Cmd> $y<-$ vector $(x)$ \# put all values in vector, a column at a time
Cmd> blk <- factor (rep(run (n), p)) \# "block number" factor
Cmd> trt <- factor (rep (run $(p), r e p(r u n(n), p))$ ) treatment numbers
Cmd> anova("y = blk + trt")
vector ( $x$ ) "unravels" data matrix $x$, column by column, into a vector of length $n-p$.
rep (run ( $n$ ) , p ) builds a vector consisting of p repetitions of each block number $1,2, \ldots, n$. rep (run (p) , rep (run (n) , p) ) builds a vector consisting of the form ( $1,1,1, \ldots, 1,2, \ldots, 2, \ldots$, $\mathrm{n}, \ldots, \mathrm{n}$ ), where there are p 1 's, p 2's, ... .

From the ANOVA you would get an $\mathrm{F}=\mathrm{MS}_{\text {treatment }} / \mathrm{MS}_{\text {error }}$ for testing $\mathrm{H}_{0}$. In a RCBD experiment, when $\mathrm{H}_{0}$ is true, F is distributed as $\mathrm{F}_{\mathrm{p}-1,(\mathrm{n}-1)(\mathrm{p}-1)}$.
However, an important component of a RCB experiment is the random assignment of the treatments to the experimental units in each block. It is what makes the use of $\mathrm{F}_{\mathrm{p}-1,(\mathrm{n}-1)(\mathrm{p}-1)}$ appropriate. In a repeated measures design this is usually impossible, even conceptually. For example, if each component is associated with a different observation time, they cannot be randomized and hence the ANOVA may not be appropriate.
Even when randomization is impossible, it may still be possible to use the ANOVA approach if you can assume an additive linear model of the following form:

$$
\mathrm{x}_{\mathrm{ij}}=\mu_{\mathrm{i}}+\mathrm{B}_{\mathrm{j}}+\varepsilon_{\mathrm{ij},} \mathrm{i}=1, \ldots, \mathrm{p}, \mathrm{j}=1, \ldots, \mathrm{n},
$$

where (i) the normal errors $\varepsilon^{\prime}$ s are independent with 0 mean; (ii) the $\varepsilon$ 's have constant variance $\sigma_{\varepsilon}{ }^{2}$; and (iii) the subject effects $\mathrm{B}_{\mathrm{j}}$ are random with expectation 0 and variance $\sigma_{\mathrm{B}}{ }^{2}$ and are independent of the $\varepsilon^{\prime}$ s. This model implies that all the correlation between different measurements on a subject comes from sharing a common subject effect.
Both the randomization model and the additive linear model imply that that the variance matrix of $\mathbf{x}$ is of the the so called intraclass correlation form

$$
\boldsymbol{\Sigma}=\left[\begin{array}{ccccc}
\sigma^{2} & \rho \sigma^{2} & \rho \sigma^{2} & \ldots & \rho \sigma^{2} \\
\rho \sigma^{2} & \sigma^{2} & \rho \sigma^{2} & \ldots & \rho \sigma^{2} \\
\rho \sigma^{2} & \rho \sigma^{2} & \sigma^{2} & \ldots & \rho \sigma^{2} \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
\rho \sigma^{2} & \rho \sigma^{2} & \rho \sigma^{2} & \ldots & \sigma^{2}
\end{array}\right]=\sigma^{2}\left[(1-\rho) \mathbf{I}_{\mathrm{p}}+\rho \mathbf{1}_{\mathrm{p}} \mathbf{1}_{\mathrm{p}}^{\prime}\right]
$$

The corresponding correlation matrix is

$$
\mathbf{R}=\left[\begin{array}{ccccc}
1 & \rho & \rho & \ldots & \rho \\
\rho & 1 & \rho & \ldots & \rho \\
\rho & \rho & 1 & \ldots & \rho \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
\rho & \rho & \rho & \ldots & 1
\end{array}\right]=(1-\rho) \mathbf{I}_{\mathrm{p}}+\rho \mathbf{1}_{\mathrm{p}} \mathbf{1}_{\mathrm{p}}
$$

For the model just given, $\sigma^{2}=\sigma_{\varepsilon}^{2}+\sigma_{B}^{2}$ and $\rho=\sigma_{B}{ }^{2} /\left(\sigma_{\varepsilon}{ }^{2}+\sigma_{B}{ }^{2}\right)$. When $\boldsymbol{\Sigma}$ is of this form, then if C is any contrast matrix, that is $\mathrm{C1}_{\mathrm{p}}=\mathbf{0}, \mathrm{V}[\mathrm{Cx}]=\mathbf{C} \Sigma \mathrm{C}^{\prime}=\sigma_{\varepsilon}{ }^{2} \mathrm{CC}^{\prime}$, where $\sigma_{\varepsilon}{ }^{2}=(1-\rho) \sigma^{2}$. This is what is exactly required for the ANOVA approach to be valid. When the ANOVA approach to testing $\mathrm{H}_{0}$ is applicable, it is preferred, since it is simpler and has greater power than $\mathrm{T}^{2}$.
Except in unlikely situations, when $\boldsymbol{\Sigma}$ is not of this form but still $\mathrm{V}[\mathbf{C x}]=\sigma_{\varepsilon}{ }^{2} \mathbf{C C}^{\prime}$, the ANOVA F-statistic is not distributed as F.

In the univariate randomized block approach, you estimate $\sigma_{\varepsilon}^{2}$ by the usual randomized block ANOVA error mean square $s^{2}=M S_{\text {error }}$ on $(n-1)(p-1)$ degrees of freedom, and
estimate $\mathrm{V}[\mathbf{C} \overline{\mathrm{x}}]$ by $\hat{\mathrm{V}}_{1}[\mathbf{C} \overline{\mathrm{x}}]=\left(\mathrm{s}^{2} / \mathrm{n}\right) \mathrm{CC}^{\prime}$. This is not the same estimate as the multivariate estimate $\hat{\mathrm{V}}[\mathbf{C} \overline{\mathbf{x}}]=\mathbf{C}(\mathbf{S} / \mathrm{n}) \mathbf{C}^{\prime}$.
It is interesting that you can express the usual ANOVA among-treatments F-statistic $\mathrm{MS}_{\text {treatment }} / \mathrm{MS}_{\text {error }}$ as $(\mathbf{C} \overline{\mathbf{x}})^{\prime}\left(\hat{V}_{1}(\mathbf{C} \overline{\mathbf{x}})\right)^{-1}(\mathbf{C} \overline{\mathbf{x}}) /(\mathrm{p}-1)$, reminiscent of the formula for $\mathrm{T}^{2}$. If $\mathrm{V}[\mathbf{C x}] \neq$ const $\times \mathrm{CC}^{\prime}$ then the ANOVA F-statistic does not have the usual F-distribution, although various approximations to its distribution are in use. See Sec. 10.17 in the MacAnova User's Guide for version 4.07.

## Two sample profile analysis

The next most complicated profile analysis situation is the two sample case.
Suppose you have two random samples $\mathbf{x}_{11}, \ldots, \mathbf{x}_{n_{1} 1}$ and $\mathbf{x}_{12}, \ldots, \mathbf{x}_{n_{2} 2}$ of sizes $n_{1}$ and $n_{2}$ from populations with means $\boldsymbol{\mu}_{1}$ and $\boldsymbol{\mu}_{2}$, respectively and common variance matrix $\boldsymbol{\Sigma}$. A typical situation might involve two treatment groups of subjects, with measurements of comparable quantities $x_{1}, x_{2}, \ldots, x_{p}$ made on each subject at times $t_{1}, t_{2}, \ldots, t_{p}$ or under $p$ different experimental conditions. You can view this situation as a two-factor repeated measures design with one within-subject factor with $p$ levels, and one between-subjects factor with 2 levels.
By analogy with a 2 by p factorial experiment, you can express the elements (individual means) $\mu_{\mathrm{i} 1}$ and $\mu_{\mathrm{i} 2}$ of $\boldsymbol{\mu}_{1}$ and $\boldsymbol{\mu}_{2}$ as

$$
\mu_{i j}=\mu+\alpha_{i}+\beta_{j}+(\alpha \beta)_{\mathrm{ij},}, i=1,2, j=1, \ldots, p .
$$

Usually the "side conditions", $\alpha_{1}+\alpha_{2}=0$ (that is, $\left.\alpha_{2}=-\alpha_{1}\right), \Sigma_{j} \beta_{\mathrm{j}}=0, \sum_{\mathrm{i}}(\alpha \beta)_{\mathrm{ij}}=(\alpha \beta)_{1 \mathrm{j}}+(\alpha \beta)_{2 \mathrm{j}}$ $=0$ (that is, $\left.(\alpha \beta)_{2 \mathrm{j}}=-(\alpha \beta)_{1 \mathrm{j}}\right)$, and $\left.\sum_{\mathrm{j}}(\alpha \beta)_{\mathrm{ij}}=0, \mathrm{i}=1,2\right)$ are assumed. These are the side conditions assumed by MacAnova for a two-way ANOVA.

The $\alpha_{i}$ are between-subjects effects, the $\beta_{j}$ 's are within-subjects effects, and the $(\alpha \beta)_{i j}$ terms are interaction effects which determine the pattern of interaction between the two factors. At each level of factor 2 , the treatment means differ by $\mu_{1 j}-\mu_{2 j}=\alpha_{1}-\alpha_{2}+(\alpha \beta)_{1 j}-(\alpha \beta)_{2 j}$. When $(\alpha \beta)_{\mathrm{ij}}=0$ for all i and j , then the elements of $\boldsymbol{\alpha}_{1}$ and $\boldsymbol{\alpha}_{2}$ are determined additively, that is, $\mu_{\mathrm{ij}}=\mu+\alpha_{\mathrm{i}}+\beta_{\mathrm{j}}$ and the differences $\mu_{1 \mathrm{j}}-\mu_{2 \mathrm{j}}=\alpha_{1}-\alpha_{2}$ do not depend on within subject level j .

## Test of parallelism (no interaction) for two samples

When $(\alpha \beta)_{\mathrm{ij}} \equiv 0$ so that there is no interaction between the two factors and $\mu_{\mathrm{ij}}=\mu+\alpha_{\mathrm{i}}+\beta_{\mathrm{j}}$, you can express each mean vector as

$$
\boldsymbol{\mu}_{\mathrm{i}}=\boldsymbol{\mu} \mathbf{1}_{\mathrm{p}}+\alpha_{\mathrm{i}} \mathbf{1}_{\mathrm{p}}+\boldsymbol{\beta}, \boldsymbol{\beta}=\left[\beta_{1}, \ldots, \beta_{\mathrm{p}}\right]^{\prime}, \mathrm{i}=1,2 .
$$

In this case, $\boldsymbol{\mu}_{1}-\boldsymbol{\mu}_{2}=\left(\alpha_{1}-\alpha_{2}\right) \mathbf{1}_{\mathrm{p}}=\left[\alpha_{1}-\alpha_{2}, \alpha_{1}-\alpha_{2}, \ldots, \alpha_{1}-\alpha_{2}\right]^{\prime}$. That is, at each level of factor 2, the two treatment means differ by a the same amount $\alpha_{1}-\alpha_{2}$. Geometrically, this means that the profiles (graphs $\mu_{1 \mathrm{j}}$ and $\mu_{2 \mathrm{j}}$ vs j ) are parallel.

When $\mathrm{H}_{0}:(\alpha \beta)_{\mathrm{ij}} \equiv 0$ is not true, that is at least one $(\alpha \beta)_{\mathrm{ij}} \neq 0$, then the profiles are not parallel and $\mathrm{C} \mu_{1} \neq \mathrm{C} \mu_{2}$.

When the profiles are parallel, $\alpha_{1}-\alpha_{2}$ is the between-subjects difference at all levels of the within-subject factor and hence it describes fully the difference between the two levels of the between-subject factor. Furthermore, the null hypothesis $\mathrm{H}_{0}: \boldsymbol{\mu}_{1}=\boldsymbol{\mu}_{2}$ reduces to the simpler hypothesis $\mathrm{H}_{0}: \alpha_{1}=\alpha_{2}$.

Similarly, let $\mathbf{c}^{\prime} \boldsymbol{\mu}_{\mathrm{i}}$ with $\mathbf{c}^{\prime} \mathbf{1}_{\mathrm{p}}=0$ define a within-subject contrast, that is, $\mathbf{c}$ defines a comparison among $\mu_{i 1}, \ldots, \mu_{i p}, i=1,2$. Then, when the assumption of parallelism is true $\left((\alpha \beta)_{i j} \equiv 0\right)$,

$$
\mathbf{c}^{\prime} \boldsymbol{\mu}_{\mathrm{i}}=\mu \mathbf{c}^{\prime} \mathbf{1}_{\mathrm{p}}+\alpha_{\mathrm{i}} \mathbf{c}^{\prime} \mathbf{1}_{\mathrm{p}}+\mathrm{c}^{\prime} \boldsymbol{\beta}=\mathrm{c}^{\prime} \boldsymbol{\beta}, \mathrm{i}=1,2
$$

does not depend on the level i of the between-group factor. This means that the $\beta_{j}$ 's fully describe the differences among the expected values at different levels of the within-subject factor. In this case, the null hypothesis that there is no difference among the means of $p$ measurements reduces to $\mathrm{H}_{0}: \beta_{1}=\beta_{2}=\ldots=\beta_{\mathrm{p}}$, that is, $\mathbf{C} \boldsymbol{\mu}_{\mathrm{i}}=\mathbf{C} \boldsymbol{\beta}=\mathbf{0}$, where $\mathbf{C}$ is a full rank $\mathrm{p}-1 \times \mathrm{p}$ matrix satisfying $\mathbf{C 1} \mathbf{1}_{\mathrm{p}}=\mathbf{0}$.
The preceding facts suggest a method of analysis to answer many of the standard questions that are asked in the univariate ANOVA of a 2 by p factorial experiment. Define the $\mathrm{n}=\mathrm{n}_{1}$ $+n_{2}$ vectors $\mathbf{w}_{k i}=\mathbf{C} x_{k i}, i=1,2, k=1, \ldots, n_{j}$ of contrasts among the elements of the $\mathbf{x}_{k i}{ }^{\prime}$ s. The $\mathbf{w}_{\mathrm{ki}}{ }^{\prime} \mathrm{s}$ have dimension $\mathrm{q}=\mathrm{p}-1$ by 1 .

## Test of parallelism (no interaction)

You can express null hypothesis $\mathrm{H}_{0}:(\alpha \beta)_{\mathrm{ij}} \equiv 0$ as $\mathrm{H}_{0}: \mathrm{E}\left[\mathbf{w}_{1}\right]=\mathrm{E}\left[\mathbf{w}_{2}\right]$, that is $\mathrm{C}_{1}=\mathrm{C} \boldsymbol{\mu}_{2}$. You can use any test of this hypothesis, in particular the two-sample Hotelling's $\mathrm{T}^{2}$ statistic based on the $\mathbf{w}_{\mathrm{ki}}{ }^{\prime} \mathrm{s}$ :

$$
\begin{aligned}
\mathrm{T}^{2} & =\left(\overline{\mathbf{w}}_{1}-\overline{\mathbf{w}}_{2}\right)^{\prime} \hat{\mathrm{V}}\left[\overline{\mathbf{w}}_{1}-\overline{\mathbf{w}}_{2}\right]^{-1}\left(\overline{\mathbf{w}}_{1}-\overline{\mathbf{w}}_{2}\right) \\
& =\left(\mathbf{C}\left(\overline{\mathbf{x}}_{1}-\overline{\mathbf{x}}_{2}\right)\right)^{\prime}\left\{\left(1 / \mathrm{n}_{1}+1 / \mathrm{n}_{2}\right) \mathrm{CSC}^{\prime}\right\}^{-1}\left(\mathrm{C}\left(\overline{\mathbf{x}}_{1}-\overline{\mathbf{x}}_{2}\right)\right) \\
& =\left\{\mathrm{n}_{1} \mathrm{n}_{2} /\left(\mathrm{n}_{1}+\mathrm{n}_{2}\right)\right\}\left(\overline{\mathbf{x}}_{1}-\overline{\mathbf{x}}_{2}\right)^{\prime} \mathbf{C}^{\prime}\left(\mathrm{CSC}^{\prime}\right)^{-1} \mathbf{C}\left(\overline{\mathbf{x}}_{1}-\overline{\mathbf{x}}_{2}\right)
\end{aligned}
$$

where $\mathbf{S}=\mathbf{S}_{\text {pooled }}=\left(n_{1}+n_{2}-2\right)^{-1}\left\{\left(n_{1}-1\right) \mathbf{S}_{1}+\left(n_{2}-1\right) \mathbf{S}_{2}\right\}$ is the pooled sample variance matrix on $\mathrm{f}_{\mathrm{e}}=\mathrm{n}_{1}+\mathrm{n}_{2}-2$ degrees of freedom. $\mathrm{T}^{2}$ has the null distribution $\left\{\mathrm{q}_{\mathrm{e}} /\left(\mathrm{f}_{\mathrm{e}}-\mathrm{q}+1\right)\right\} \mathrm{F}\left(\mathrm{q}, \mathrm{f}_{\mathrm{e}}-\mathrm{q}+1\right)=$ $\left\{(\mathrm{p}-1)\left(\mathrm{n}_{1}+\mathrm{n}_{2}-2\right) /\left(\mathrm{n}_{1}+\mathrm{n}_{2}-\mathrm{p}\right)\right\} \mathrm{F}_{\mathrm{p}-1, \mathrm{n}_{1}+\mathrm{n}_{2}-\mathrm{p}}$ because the dimension $\mathbf{w}$ is $\mathrm{q}=\mathrm{p}-1$. As before, the value of $T^{2}$ does not depend on the particular $C$ chosen. You would reject $H_{0}$ if $\mathrm{T}^{2}>\left\{(\mathrm{p}-1)\left(\mathrm{n}_{1}-\mathrm{n}_{2}+2\right) /\left(\mathrm{n}_{1}+\mathrm{n}_{2}-\mathrm{p}\right)\right\} \mathrm{F}_{\mathrm{p}-1, \mathrm{n}_{1}+\mathrm{n}_{2}-\mathrm{p}}(\alpha)$.

## Two sample test of within-subject equality of means, assuming parallelism.

When there is no interaction, that is, when the profiles are parallel, the $n=n_{1}+n_{2}$ vectors $\mathbf{w}_{11}, \ldots, \mathbf{w}_{\mathrm{n}_{1} 1}, \mathbf{w}_{12}, \ldots, \mathbf{w}_{\mathrm{n}_{2} 2}$ all have common mean vector $\mathbf{C} \boldsymbol{\beta}$. Since the null hypothesis of no within-subject differences, $\left(\mathrm{H}_{0}: \beta_{1}=\beta_{2}=\ldots=\beta_{\mathrm{p}}\right)$ is equivalent to $\mathrm{H}_{0}: \mathrm{E}\left[\mathbf{w}_{\mathrm{ij}}\right]=\mathbf{C} \boldsymbol{\beta}=\mathbf{0}$, you can test it by a single sample Hotelling's $T^{2}$, treating all $n_{1}+n_{2} \mathbf{w}_{\mathrm{ki}}$ 's as a sample from a single population with sample mean $\overline{\mathbf{w}}=\left(\mathrm{n}_{1} \overline{\mathbf{w}}_{1}+\mathrm{n}_{2} \overline{\mathbf{w}}_{2}\right) /\left(\mathrm{n}_{1}+\mathrm{n}_{2}\right)=\mathbf{C} \overline{\mathbf{x}}$, where $\overline{\mathbf{x}}=$ $\left(n_{1} \overline{\mathbf{x}}_{1}+\mathrm{n}_{2} \overline{\mathbf{x}}_{2}\right) /\left(\mathrm{n}_{1}+\mathrm{n}_{2}\right)$. The estimated variance matrix of $\overline{\mathbf{x}}$ is $\hat{\mathrm{V}}[\overline{\mathbf{x}}]=(1 / \mathrm{n}) \mathbf{S}$ and therefore the estimated variance matrix of $\overline{\mathbf{w}}=\mathbf{C} \overline{\mathbf{x}}$ is $\hat{\mathrm{V}}[\overline{\mathbf{w}}]=\hat{\mathrm{V}}[\mathbf{C} \overline{\mathbf{x}}]=(1 / \mathrm{n}) \mathrm{CSC}^{\prime}$. Hotelling Æs $\mathrm{T}^{2}$ is

$$
\mathrm{T}^{2}=\overline{\mathbf{w}}^{\prime} \hat{\mathrm{V}}[\overline{\mathbf{w}}]^{-1} \overline{\mathbf{w}}=(\mathbf{C} \overline{\mathbf{x}})^{\prime}\left\{(1 / \mathrm{n}) \mathrm{CSC}^{\prime}\right\}^{-1} \mathbf{C} \overline{\mathbf{x}}
$$

with null distribution $\left\{(\mathrm{p}-1)\left(\mathrm{n}_{1}+\mathrm{n}_{2}-2\right) /\left(\mathrm{n}_{1}+\mathrm{n}_{2}-\mathrm{p}-1\right)\right\} \mathrm{F}_{\mathrm{p}-1, \mathrm{n}_{1}+\mathrm{n}_{2}-\mathrm{p}-1}$. This provides a test of the "main effects" of the within-subject factor, assuming no interaction.
If the null hypothesis of parallelism is rejected, it is certainly evidence that $C \overline{\boldsymbol{\mu}}_{1} \neq \mathrm{C} \overline{\boldsymbol{\mu}}_{2}$ and that either $\mathbf{C} \overline{\boldsymbol{\mu}}_{1} \neq 0$ or $\mathbf{C} \overline{\boldsymbol{\mu}}_{2} \neq 0$ or both.

## Test of equality of the two group mean vectors, assuming parallelism

Again assume that the effects of within-subject treatment and between-subject treatment are additive, that is, that the the two profiles are parallel. Then you may want to test the null hypothesis that the two group mean vectors are identical, that is, the hypothesis that there is no non-zero between-subject factor effect. Symbolically this is $H_{0}: \alpha_{1}=\alpha_{2}=0$. You could test $\mathrm{H}_{0}$ by an ordinary two sample Hotelling's $\mathrm{T}^{2}$ test of $\mathrm{H}_{0}: \boldsymbol{\mu}_{1}=\boldsymbol{\mu}_{2}$, but this does not take advantage of the assumption that the profiles are parallel. With this assumption, the problem can reduced to a univariate problem.

Let $\boldsymbol{\mathbb { 0 }}$ be any fixed $p$ by 1 vector that defines a linear combination of the $p$ responses such that $\boldsymbol{\mathbb { Q }} \mathbf{1}_{\mathrm{p}}=1$, for example $\boldsymbol{\mathbb { }}=\mathbf{1}_{\mathrm{p}} / \mathrm{p}$. Then, assuming the additive model $\mathrm{E}\left[\mathbf{x}_{\mathrm{i}}\right]=\boldsymbol{\mu}_{\mathrm{i}}=\mu \mathbf{1}_{\mathrm{p}}+$ $\alpha_{i} 1_{p}+\beta$,

$$
\mathrm{E}\left[\boldsymbol{0}^{\prime} x_{\mathrm{ij}}\right]=\boldsymbol{0}^{\prime} \boldsymbol{\mu}_{\mathrm{j}}=\mu^{\prime} \boldsymbol{\mathbf { l }}^{\prime} \mathbf{1}_{\mathrm{p}}+\alpha_{\mathrm{j}} \mathbf{0}^{\prime} \mathbf{1}_{\mathrm{p}}+\boldsymbol{\mathbf { 0 }} \boldsymbol{\beta}=\mu+\alpha_{\mathrm{j}}, \mathrm{i}=1, \ldots, \mathrm{n}_{\mathrm{j}}, \mathrm{j}=1,2
$$

 equality with a univariate two sample t-statistic

$$
\begin{aligned}
& \mathrm{t}=\left(\boldsymbol{\underline { 0 }}^{\prime} \overline{\mathbf{x}}_{1}-\boldsymbol{\underline { 0 }}^{\prime} \overline{\mathbf{x}}_{2}\right) / \widehat{\mathrm{SE}}\left[\boldsymbol{0}^{\prime} \overline{\mathbf{x}}_{1}-\boldsymbol{\underline { 0 }}^{\prime} \overline{\mathbf{x}}_{2}\right]= \\
& \left(\boldsymbol{0}^{\prime} \overline{\mathbf{x}}_{1}-\boldsymbol{\underline { 0 }}{ }^{\prime} \overline{\mathbf{x}}_{2} /\left\{\left(1 / \mathrm{n}_{1}+1 / \mathrm{n}_{2}\right) \boldsymbol{\boldsymbol { 0 }}{ }^{\prime} \boldsymbol{S} \boldsymbol{\boldsymbol { 0 }}\right\}^{1 / 2} .\right.
\end{aligned}
$$

When $H_{0}$ is true, $t$ is distributed as Student's $t_{f_{e}}=t_{n_{1}+n_{2}-2}$. Unfortunately, there is no unique choice for $\boldsymbol{\emptyset}$. The conventional choice is to weight all responses equally, that is, to put $\mathbf{Q}=$ $\mathrm{p}^{-1} \mathbf{1}_{\mathrm{p}}$, so that $\boldsymbol{\underline { D }}^{\prime} \mathbf{x}=\mathrm{p}^{-1} \Sigma_{1 \leq \mathrm{j} \leq \mathrm{p}} \mathrm{x}_{\mathrm{j}}$ is the average response across the levels of the within-subject factor. Ideally you would like to choose $\boldsymbol{0}$ such that $\boldsymbol{\underline { 0 }} \boldsymbol{\Sigma} \boldsymbol{\underline { 0 }} /\|\boldsymbol{0}\|^{2}$ is minimized, since this will maximize the power of the $t$-test.

## Profile Analysis

## Multi-sample profile analysis

The concepts involved in two sample profile analysis are easily extended to several samples. Now you analyze $g$ random samples $\mathbf{x}_{11}, \ldots, \mathbf{x}_{n_{1}, 1}, \mathbf{x}_{12}, \ldots, \mathbf{x}_{n_{2}, 2}, \ldots, \mathbf{x}_{1 g}, \ldots, \mathbf{x}_{\mathrm{n}_{\mathrm{g}} \mathrm{g}}$ from $g$ populations with means $\boldsymbol{\mu}_{1}, \boldsymbol{\mu}_{2}, \ldots, \boldsymbol{\mu}_{\mathrm{g}}$ respectively and common variance matrix $\boldsymbol{\Sigma}$. A typical situation might involve $g$ treatment groups of subjects, with measurements of comparable quantities $x_{1}, x_{2}, \ldots, x_{p}$ made on each subject at times $t_{1}, t_{2}, \ldots, t_{p}$ or under $p$ different experimental conditions. This situation may be viewed as a two-factor repeated measures design with one within-subject factor with p levels, and 1 between-subjects factor with $g$ levels.
Precisely as in the two sample case, you can decompose each mean as

$$
\mu_{i j}=\mu+\alpha_{i}+\beta_{j}+(\alpha \beta)_{i j}, i=1, \ldots, g, j=1, \ldots, p
$$

The $\alpha_{i}^{\prime}$ 's are the between-groups (between subjects) effects, the $\beta_{j}$ 's are the betweenmeasurements within-subjects effects, and the $(\alpha \beta)_{\mathrm{ij}}{ }^{\prime} \mathrm{s}$ are interaction effects expressing how the within-subjects effects differ among groups. Usually it is assumed that $\sum_{i} \alpha_{i}=\sum_{j} \beta_{j}=$ $\sum_{\mathrm{i}}(\alpha \beta)_{\mathrm{ij}}=\sum_{\mathrm{j}}(\alpha \beta)_{\mathrm{ij}}=0$.

## Test of parallelism with several groups

If all the interaction terms are zero, then for each group, the differences $\mu_{\mathrm{ij}_{1}}-\mu_{\mathrm{ij}_{2}}=\beta_{\mathrm{j}_{1}}-\beta_{\mathrm{j}_{2}}$ of the components $\mu_{\mathrm{ij}}$ of $\boldsymbol{\mu}_{\mathrm{i}}$ are the same for all groups. Geometrically this means that the g mean profiles are parallel. For this reason a test of $\mathrm{H}_{0}:(\alpha \beta)_{i j}=0, i=1, \ldots, g, j=1, \ldots, p$ is often called a test for parallelism. If $\mathrm{H}_{0}$ is true, then, as with the case of two groups,

$$
\boldsymbol{\mu}_{\mathrm{i}}=\boldsymbol{\mu} \mathbf{1}_{\mathrm{p}}+\alpha_{\mathrm{i}} \mathbf{1}_{\mathrm{p}}+\boldsymbol{\beta}, \text { where } \boldsymbol{\beta}=\left[\beta_{1}, \beta_{2}, \ldots, \beta_{\mathrm{p}}\right]^{\prime}
$$

Let C be a rank $\mathrm{p}-1$ contrast matrix, that is $\mathrm{C}^{\prime} \mathbf{1}_{\mathrm{p}}=\mathbf{0}_{\mathrm{p}-1}$. Under the hypothesis of parallelism, $\mathbf{C} \boldsymbol{\mu}_{\mathrm{i}}=\mathbf{C} \boldsymbol{\beta}, \mathrm{i}=1, \ldots, \mathrm{~g}$. That is, if $\mathbf{w}=\mathbf{C x}$, the means $\boldsymbol{\mu}_{\mathrm{w}, \mathrm{i}}=\mathbf{C} \boldsymbol{\mu}_{\mathrm{i}}=\mathbf{C} \boldsymbol{\beta}$ are identical for all i . Therefore any test the hypothesis $\boldsymbol{\mu}_{\mathrm{w}, 1}=\boldsymbol{\mu}_{\mathrm{w}, 2}=\ldots=\boldsymbol{\mu}_{\mathrm{w}, \mathrm{g}}$ provides a test of parallelism. You can do this by a p-1 dimensional multivariate analysis of variance MANOVA based on $\mathbf{w}_{\mathrm{ij}}=$ $C x_{i j}, i=1, \ldots, n_{j}, j=1, \ldots, g$. If $\mathbf{H}$ and $E$ are the $p$ by $p$ hypothesis and error matrices in $a$ MANOVA based on $\mathbf{x}$, the $\mathrm{p}-1$ by $\mathrm{p}-1$ matrices $\mathbf{H}_{\mathrm{w}}=\mathbf{C H C}$ and $\mathbf{E}_{\mathrm{w}}=\mathbf{C E C}$, are the hypothesis and error matrices for the analysis of $\mathbf{w}$. You can use any of the usual MANOVA tests based on $\mathbf{H}_{w}$ and $\mathbf{E}_{\mathrm{w}}$.
If the profiles are not parallel, that is, there is interaction between the among-group factor and the within-subject factor, it usually means that there is little point in testing "main effects", since the presence of interaction means that the effects of each factor depend on the level of the other factor. Hence, to examine main effects, you usually assume parallelism.

## Multi sample test of within-subject equality of means, assuming parallelism

Assuming parallelism (no interaction), all the variable means within a subject are equal whenever $\beta_{1}=\beta_{2}=\ldots=\beta_{p}=\beta$, that is $\boldsymbol{\beta}=\beta 1_{p}$. When $\boldsymbol{\beta}=\beta 1_{p}$, and only $\boldsymbol{\beta}=\beta \boldsymbol{1}_{p}$, then $\mathrm{C} \boldsymbol{\beta}=$ $\beta \mathbf{C 1} 1_{\mathrm{p}}=\mathbf{0}$. Thus you can test $\mathrm{H}_{0}: \beta_{1}=\beta_{2}=\ldots=\beta_{\mathrm{p}}$ by testing $\boldsymbol{\mu}_{\mathrm{w}}=\mathrm{E}[\mathbf{C x}]=\mathbf{0}$. You can do this using Hotellings $T^{2}=\overline{\mathbf{w}}^{\prime} \hat{V}[\overline{\mathbf{w}}]^{-1} \overline{\mathbf{w}}$, where $\overline{\mathbf{w}}=\mathbf{C} \overline{\mathbf{x}}, \overline{\mathbf{x}}=\left(\sum_{\mathrm{i}} \mathrm{n}_{\mathrm{i}} \overline{\mathbf{x}}_{\mathrm{i}}\right) / \mathrm{n}=\sum_{j} \sum_{\mathrm{i}} \mathbf{x}_{\mathrm{ij}} / \mathrm{n}$ is the grand
mean vector, and $\hat{V}[\overline{\mathbf{w}}]=n^{-1} \mathbf{C S} \mathbf{C}^{\prime}$. Here $\mathrm{n}=\sum_{i} \mathrm{n}_{\mathrm{i}}$ is the total number of cases (subjects) and $\mathbf{S}_{\mathrm{p}}=\mathbf{E} /(\mathrm{n}-\mathrm{g})$ is the pooled covariance matrix on $\mathrm{f}_{\mathrm{e}}=\mathrm{n}-\mathrm{g}$ degrees of freedom. Assuming normality and equality of covariance matrices $\left(\boldsymbol{\Sigma}_{1}=\boldsymbol{\Sigma}_{2}=\ldots=\boldsymbol{\Sigma}_{\mathrm{p}}\right)$, under $\mathrm{H}_{0}$, since the dimension of $\mathbf{w}$ is $\mathrm{q}=\mathrm{p}-1, \mathrm{~T}^{2}$ is distributed as

$$
\left\{\mathrm{f}_{\mathrm{e}}-\mathrm{q} /\left(\mathrm{f}_{\mathrm{e}}-\mathrm{q}-1\right)\right\} \mathrm{F}_{\mathrm{q}, \mathrm{f}_{\mathrm{e}}-\mathrm{q}-1}=\{(\mathrm{n}-\mathrm{g})(\mathrm{p}-1) /(\mathrm{n}-\mathrm{g}-\mathrm{p})\} \mathrm{F}_{\mathrm{p}-1, \mathrm{n}-\mathrm{g}-\mathrm{p}} .
$$

## Multi-sample test of equality of group mean vectors, assuming parallelism

Now suppose that you want to test the null hypothesis $H_{0}: \boldsymbol{\mu}_{1}=\boldsymbol{\mu}_{2}=\ldots=\boldsymbol{\mu}_{\mathrm{g}}$, assuming that the mean vectors are of the form $\boldsymbol{\mu}_{\mathrm{i}}=\mu \mathbf{1}_{\mathrm{p}}+\alpha_{\mathrm{i}} \mathbf{1}_{\mathrm{p}}+\boldsymbol{\beta}$. Clearly, when you can assume parallelism, $\mathrm{H}_{0}$ is true if and only if $\alpha_{1}=\alpha_{2}=\ldots=\alpha_{\mathrm{g}}$. As in the two sample case, let $\boldsymbol{\mathbb { 0 }}$ be a p by 1 vector such that $\boldsymbol{Q}^{\prime} \boldsymbol{1}_{\mathrm{p}}=1$. Then, if $\gamma_{\mathrm{i}}=\boldsymbol{\Omega}^{\prime} \boldsymbol{\mu}_{\mathrm{i}}, \gamma_{\mathrm{i}}=(\mu+\beta)+\alpha_{\mathrm{i}}$, with $\beta=\boldsymbol{0}$ ' $\boldsymbol{\beta}$. Then $\alpha_{1}=\alpha_{2}$ $=\ldots=\alpha_{\mathrm{g}}$ if and only if $\gamma_{1}=\gamma_{2}=\ldots=\gamma_{\mathrm{g}}$. But, if $\mathrm{y} \equiv \boldsymbol{\mathbf { 0 }}{ }^{\prime} \mathbf{x}, \mathrm{y}$ is a univariate random variable whose mean in group $i$ is $\gamma_{i}$. The usual test in the univariate situation is a univariate oneway ANOVA F-statistic on $g-1$ and $f_{e}=n-g$ degrees of freedom. If $\mathbf{H}$ and $\mathbf{E}$ are the hypothesis and error matrices for a one-way MANOVA of the hypothesis $\mathrm{H}_{0}: \boldsymbol{\mu}_{1}=\boldsymbol{\mu}_{2}=\ldots=$ $\boldsymbol{\mu}_{\mathrm{g}}$, then the numerator and denominator sums of squares in the F-statistic are $\mathbf{\underline { 0 }} \mathbf{H} \mathbf{0}$ and $\boldsymbol{0}{ }^{\prime}$ E $\boldsymbol{0}$, respectively. As with the two sample case, the conventional choice for $\boldsymbol{0}$ is $\boldsymbol{0}=$ $(1 / p) 1_{p}$, and the analysis reduces to an ANOVA on the equally weighted average of all $p$
 $\sum_{\mathrm{i}} \sum_{\mathrm{j}} \mathrm{e}_{\mathrm{ij}} / \mathrm{p}^{2}$, where $\mathbf{H}=\left[\mathrm{h}_{\mathrm{ij}}\right]$ and $\mathbf{E}=\left[\mathrm{e}_{\mathrm{ij}}\right]$.

