Regression with circular data

We follow the methods proposed by Jolien Cremers and collaborators.

Basic idea
We have $i = 1, ..., n$ individuals, each with multiple observations $j = 1, ...J_i$, for a total number of observations equal to $\sum_i J_i = N$. The observations are angles, which we call θ_{ij} and measure in radians, so that $-\pi \leq \theta_{ij} \leq \pi$.
The jth angle for the ith person can be represented as a two-dimensional unit vector in R^2, which we denote u_{ij}, where:

$$u_{ij} = \frac{y_{ij}}{r_{ij}},$$

with $r_{ij} = ||y_{ij}||$, the length of the unobservable bivariate vector y_{ij} and $u_{ij} = (\cos\theta_{ij}, \sin\theta_{ij})$. See Fig. 1.

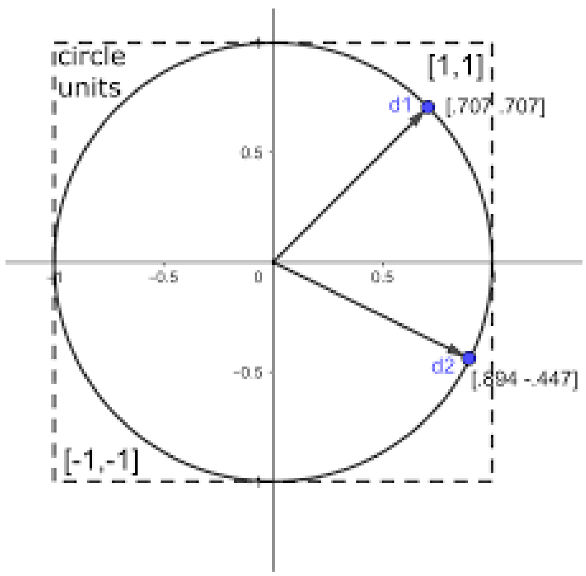

Figure 1: Unit vector

Regression
Suppose that for each participant we also have information on independent regressors. In our case, for each person we have age, gender and dominant writing hand. Age is recorded as a categorical variable, with two levels: age between 18 and 40 years, or age 41 years and older.
We define three dummy variables: x_1 do denote age group, x_2 to denote gender, and x_3 to denote dominant hand. The base categories (for which the dummy variables take on the value 0) are: age 18-40, female and left-handed.
We fit two regression models, one to each of the elements of $y_{ij} = (y^I_{ij}, y^{II}_{ij})$, so that

$$y^I_{ij} = \beta^I_0 + \beta^I_1 x_{1i} + \beta^I_2 x_{2i} + \beta^I_3 x_{3i} + \alpha^I_i + \epsilon^I_{ij}$$

$$y_{ij}^{II} = \beta_0^{II} + \beta_1^{II} x_{1i} + \beta_2^{II} x_{2i} + \beta_3^{II} x_{3i} + \alpha_i^{II} + \epsilon_{ij}^{II},$$

where the $(\alpha_i^I, \alpha_i^{II})$ are random intercepts with mean 0 and constant variance for each participant, accounting for the fact that we have repeated observations for each person, and the residuals are normally distributed.
There are eight possible combinations of the fixed effects, so expressions for the predicted values for (y_{ij}^I, y_{ij}^{II}) are shown below.

Category	$\hat{y}^I$	$\hat{y}^{II}$
Female, 18-40, Left	b_0^I	b_0^{II}
Female, 18-40,Right	$b_0^I + b_3^I$	$b_0^{II} + b_3^{II}$
Female, 41+, Left	$b_0^I + b_1^I$	$b_0^{II} + b_1^{II}$
Female, 41+, Right	$b_0^I + b_1^I + b_3^I$	$b_0^{II} + b_1^{II} + b_3^{II}$
Male, 18-40, Left	$b_0^I + b_2^I$	$b_0^{II} + b_2^{II}$
Male, 18-40, Right	$b_0^I + b_2^I + b_3^I$	$b_0^{II} + b_2^{II} + b_3^{II}$
Male, 41+, Left	$b_0^I + b_1^I + b_2^I$	$b_0^{II} + b_1^{II} + b_2^{II}$
Male, 41+, Right	$b_0^I + b_1^I + b_2^I + b_3^I$	$b_0^{II} + b_1^{II} + b_2^{II} + b_3^{II}$

Table 1: Estimated values of the two linear components for each type of person. The b denote estimated regression coefficients.

Circular predicted values

Once we have computed, for each person, a pair of predicted values $(\hat{y}_{ij}^I, \hat{y}_{ij}^{II})$ using the expressions in Table 1, we can convert those into circular predicted values $\hat{\theta}_{ij}$ as follows:

$$\begin{aligned}
\text{If } \hat{y}^I > 0 \qquad & \hat{\theta} = \arctan\left(\frac{\hat{y}^{II}}{\hat{y}^I}\right) \\
\text{If } \hat{y}^I \leq 0, \hat{y}^{II} \geq 0 \qquad & \hat{\theta} = \arctan\left(\frac{\hat{y}^{II}}{\hat{y}^I}\right) + \pi \\
\text{If } \hat{y}^I < 0, \hat{y}^{II} < 0 \qquad & \hat{\theta} = \arctan\left(\frac{\hat{y}^{II}}{\hat{y}^I}\right) - \pi \\
\text{If } \hat{y}^I = 0, \hat{y}^{II} > 0 \qquad & \hat{\theta} = \frac{\pi}{2} \\
\text{If } \hat{y}^I = 0, \hat{y}^{II} < 0 \qquad & \hat{\theta} = -\frac{\pi}{2}.
\end{aligned}$$

If both $\hat{y}^I = 0, \hat{y}^{II} = 0$ then $\hat{\theta}$ is not defined.

Results