Statistical methods for the comparison of vestibular response vectors: post-hoc testing for phase differences. Scott H Seidman and Jeremy Sharib. Univ of Rochester, Depts. Biomedical Engineering and Neurobiology&Anatomy, NY 168.14

Introduction

Data from vestibular experiments are often expressed as vectors in a polar coordinate representation, requiring both a magnitude and phase to characterize a response. Rigorous statistical comparison of response distributions is facilitated by a transformation of response vectors into a Cartesian representation in complex space, followed by multivariate statistical testing techniques. It is often necessary, however, to attribute any significant differences that arise to changes in magnitude and phase, and the Cartesian analysis is not conducive to such determinations. Post-hoc testing for changes in phase is not straightforward, largely because the distribution of response phase does not necessarily reflect the vectorial mean response.

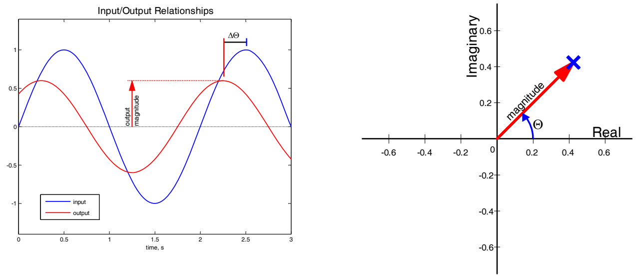
A convenient approach for post-hoc phase comparison is presented. By the proposed method, the phases of two vector populations are considered to be different if the two populations are not collinear with the origin. Determination of collinearity is assessed using standard dummy variable techniques. Data from the two populations to be compared are fit to a zero-intercept linear model in Cartesian space (i.e., a line forced through the origin), with a regressor representing the difference in slope between the two populations. A statistically significant non-zero value of this regressor is thus indicative of a significant phase difference.

Methods

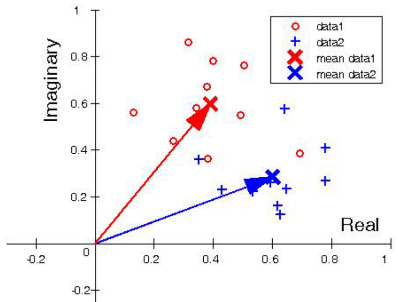
- Simulated data sets (N=10 and N=50) were statistically compared using two methods: a comparison of slope in the complex plane using **a zero-intercept multivariate regression** with dummy variables on the vectorial data, and a direct t-test using only the phase angles.
- Minimum detectable phase difference, magnitude = 1.
    - The first data set had a magnitude of 1 and phase of 45°. The second set had an magnitude of 1 and an angle differing from the first set by pi/128 radian increments (1.4°) to a maximum of 35.2°. Standard deviations of both sets were equal, and ranged from 0.05 in the complex plane (2.9° when magnitude = 1) to 0.3 (16.7° when magnitude = 1), in increments of 0.05. A total of 150 data sets were generated (25 phase angles x 6 standard deviations). P-values were corrected for multiple comparison using Holm's Procedure.
    - The minimum detectable phase difference for each standard deviation was recorded for each of the statistical methods, and compared.
- Minimum magnitude for 22.5° phase difference.
    - The first simulated data set was magnitude 1, phase 45°. The second set ranged from magnitude 0.05 to magnitude 1 in increments of 0.05. Standard deviations of both sets were the same, and 0.05 or 0.15 in the complex plane (20 magnitudes x 2 standard deviations = 40 comparisons). Resulting p-values were adjusted using Holm's procedure.
    - The minimum magnitude of the second data set for which a significant difference was found using each statistical method was noted and compared.

Representation of Data on Complex Plane

1



2 Sample Data Sets

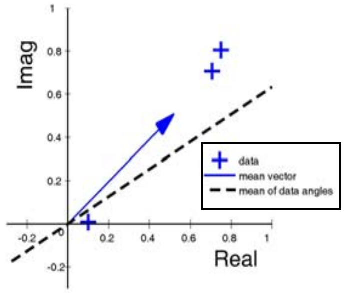


- Data sets are best compared using Multivariate techniques on complex representation.
- This does not resolve the problem of the comparison of the phase of two data sets

Problems Inherent to Post-hoc Phase Testing

- Phase is periodic, and wrapping issues can cause statistical problems (e.g., is there a 90° lead, or a 270° lag??).
- Standard deviations in phase are related to standard deviations in the complex plane by a tangent relationship.
- The mean phase of the data points does not necessarily correspond to the vector mean of the data, thus phase-only comparison might not reflect the vectorial distribution of the data.

3



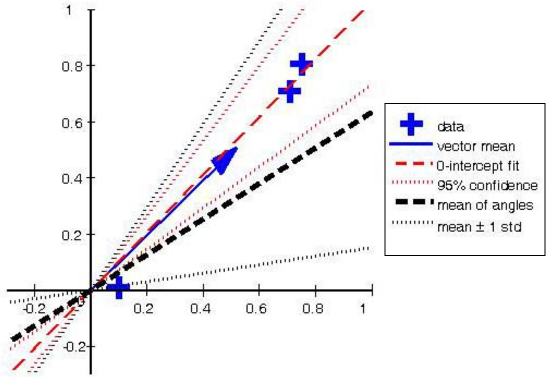
Proposed Solution: Compare Slopes of Zero-Intercept Regression

$$y = \beta x + \varepsilon$$

$$S(\beta) = \sum_{i=1}^{n} (y_i - \beta x_i)^2$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2} \qquad \hat{\beta} - t_{\alpha/2,n-1}\sqrt{\frac{MS_E}{\sum x_i^2}} \le \beta \le \hat{\beta} + t_{\alpha/2,n-1}\sqrt{\frac{MS_E}{\sum x_i^2}}$$

4 Utility of zero-intercept model



- Zero-intercept fit somewhat analogous to vector notation
- Fit is automatically weighted to larger vectors
- Slope of fit approaches the phase of the vector mean
- No wrapping problems
- Formal expressions for confidence interval on slope facilitates phase comparison

Multivariate Implementation with Dummy Variables

The catenated data matrix:

$$\begin{array}{ccc} y_{11} & x_{11} & 0 \\ y_{12} & x_{12} & 0 \\ . & . & . \\ y_{1n} & x_{1n} & 0 \\ y_{21} & x_{21} & 1 \\ y_{22} & x_{22} & 1 \\ . & . & . \\ y_{2m} & x_{2m} & 1 \end{array}$$

Is fit to the model:

$$y = \beta_1 x + \beta_2 x D$$

$$when \quad D = 0$$

$$y = \beta_1 x$$

$$when \quad D = 1$$

$$y = (\beta_1 + \beta_2)x$$

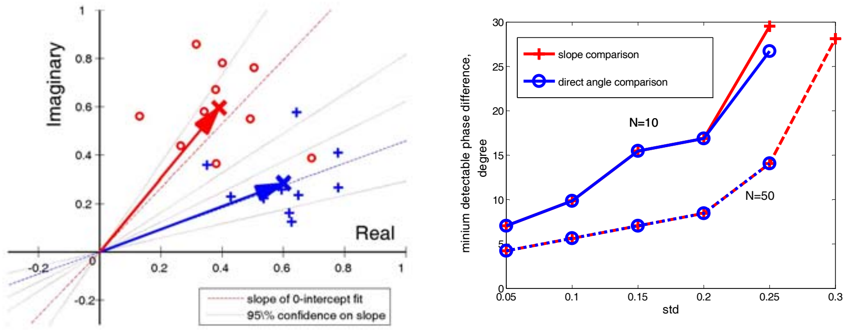
Thus, a significant value of β_2 indicates a difference in slope between the two sets

Steps to test for phase difference:

1. Manova to determine difference
2. Test to see if either data set includes origin in complex plane, which would make phase meaningless.
3. Test for difference in slope of zero-intercept fit.
4. Check non-significant difference for 180° phase shift.

Results

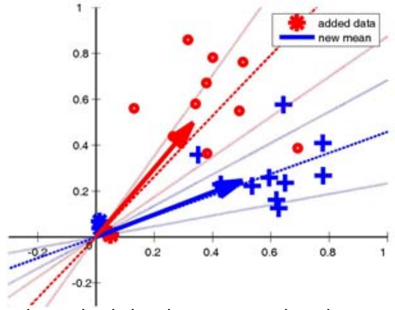
5



When data were distributed normally in the complex plane:

- no differences between the slope-comparison method and the simple t-test were noted for N=10
- For N=50, STD of 0.3, the slope method was able to distinguish a phase difference of 33.7 degrees, while the simple t-test was unable to distinguish any difference for any of our simulations at that high STD.
- A 22.5° phase difference between the comparison set and a unit vector was detectable for magnitudes as small as 0.35 when STD was 0.05 (8.2° at that magnitude) using the slope comparison, and 0.3 when using a t-test directly on phase. For STD=0.15, the slope method detected a phase change at magnitude=0.7, while phase comparison detected the difference at 0.55.

6



Angle(mean vector)	Mean(Angles)
26.9	33.7
56.3	49.3

The above simulation demonstrates the robustness of the slope method in the presence of low-magnitude phase outliers, which have little impact upon the phase of the mean vector, while pulling the means of the angles closer together. These two sets are not significantly different using the direct t-test, but are easily distinguishable via the slope method.

Conclusions

Comparisons of slopes of zero-intercept fits yield little difference from direct phase comparisons when data are distributed normally in the complex plane

When data are not distributed normally, the slope comparison method is more robust in the presence of response vectors that are small compared to the mean (when phase measurements tend to be most suspect) but by its nature is more sensitive to large-magnitude outliers.

Calkins, DS, Examination of two methods for statistical analysis of data with magnitude and direction emphasizing vestibular research applications. J Vestibular Res 8(1998)334-340

Myers, RH, Classical and Modern Regression with Applications, PWS-Kent, Boston, 1990

Supp by NIH DC04153