

# projection analysis

mike freund

May 4, 2020

## intro

See bishop (p 182) for projection onto hyperplane, Izenman (p 242, eq 8.9) for expression of linear discriminant rule. (also p 107–119 HTF.)

## notation

- vertex  $v$  in  $1, \dots, V$
- TR  $t$  in  $1, \dots, T$
- $\mathbf{m}_{(condition)}$ : class centroid (beta pattern vector) of length  $V$ , *condition* in *incon.*, *congr.*
- $\cdot'$ : from test set
- $\mathbf{e}_t$ : residual vector of length  $V$ , at TR =  $t$
- $\mathbf{S}^*$ : regularized covariance matrix of residuals ( $V \times V$ ; asterisk indicates regularization)

## within-run version

First validate the method by examining the distribution about the discriminant *within each run*. This is perhaps the most ‘liberal’ test as it does not require the discriminant to be aligned across scanning runs.

Select a condition to represent the positive end of the discriminant. Calculate the difference between class centroids and prewhiten. This forms the discriminant function,  $\mathbf{b}$ .

$$\mathbf{b} = (\mathbf{m}_{incon.} - \mathbf{m}_{congr.})\mathbf{S}^{*-1}$$

A residual is centered at the mean pattern and projected onto the discriminant.

$$\hat{e}_t = \mathbf{b}(\mathbf{e}_t - \bar{\mathbf{m}})$$

This will equal zero when  $\mathbf{e}_t$  is equidistant between class means.

However the scale will depend on the length of the discriminant function. Scaling this expression by the length of the discriminant function will yield the distance to hyperplane. Scaling by the square of the length will render the units a ratio between distance to hyperplane relative to distance between class mean.

$$\hat{e}_t = \frac{\mathbf{b}(\mathbf{e}_t - \bar{\mathbf{m}})}{\|\mathbf{b}\|_2^2}$$

E.g.,  $\hat{e}_t = 1$  would indicate that  $\mathbf{e}_t$  was the same distance from hyperplane as the class centroid,  $\mathbf{m}_{incon.}$ .

## validation: positive / negative control analyses

Detrend BOLD timeseries with baseline model of regression. Project detrended timeseries from ROIs and non-ROIs onto discriminant. Regress projection onto events design matrix.<sup>1</sup> Examine this projection to validate method.

- Was trial info recovered?  $\beta_{(incon.)} > 0$  and  $\beta_{(congr.)} < 0$ . Generalization across item types (PC50, Bias)?
- Sanity check: plot projection distributions per subject. Hopefully bimodal. Prelim look at inter-trial intervals — where do these TRs fall on discriminant?
- Cosine similarity between recovered  $\beta$  and  $\mathbf{b}$ . Hopefully positive.
- In ROIs, at post-trial TRs, do residuals move along discriminant as function of previous trial type? Does this effect disappear in non-ROIs?
- some form of autocorrelation analysis: how bad is problem?

Depending on success, these projections may be used to address the central question.

## cross-validated version

Next establish robustness to scanning run by cross-validating the discriminant *across runs*.

This can either be done by assuming the same threshold (zero-point along the discriminant function)

$$\hat{e}'_t = \mathbf{b}(\mathbf{e}'_t - \bar{\mathbf{m}}) / \|\mathbf{b}\|_2^2$$

or by using the threshold estimated in the test run

$$\hat{e}''_t = \mathbf{b}(\mathbf{e}'_t - \bar{\mathbf{m}}') / \|\mathbf{b}\|_2^2$$

Using the test run threshold would be preferable if large, systematic (i.e., across all TRs) shifts in the mean pattern are expected between runs.

The adequacy of either threshold can be checked empirically: does the boundary drawn give good separation between projections? If not, perhaps an optimal threshold can be selected.

## validation: positive / negative control analyses

- Repeat same validation analyses as above.
- Calculate cosine angle between discriminants across runs. Hopefully positive. (Can assess impact of prewhitening, too.)

## analyses of theoretic interest

---

<sup>1</sup>Hierarchical regression *may* be appropriate, if sufficiently complex model can be fit. E.g., could collapse all  $n$  trial type knots into single regressors, deviation coded, such that *incon.*  $> 0$  and *congr.*  $< 0$ . This would reduce to  $n$  df per subject (estimated in random effects). Perhaps not all of those need to be estimated.