Bayesian inference for correlations between brain activity patters

Jörn Diedrichsen (jdiedric@uwo.ca)
Western Institute of Neuroscience
London, Ontario, Canada

Mahdiyar Shahbazi (mshahba9@uwo.ca)
Western Institute of Neuroscience
London, Ontario, Canada

Abstract

The relationship between activity patterns in response to different conditions provides important insights into the computations occurring in a brain area. To fully characterize the representational geometry, it is often desirable to establish the correlation between two activity patterns (or between two representational hyper-planes), independent of the size of the activation. Traditional point-estimates of correlation coefficients between patterns are biased and not suited for inference. This is especially true for functional magnetic resonance imaging (fMRI) data, which is corrupted by substantial measurement noise. Here we propose a Bayesian approach, which approximates the posterior distribution of the correlation coefficient. This approach allows valid inferences, both when comparing a correlation coefficient against a fixed value (one-sample problem), as well as comparing two correlation coefficients across two different regions or groups of subjects (two-sample problem). The utility of the approach is demonstrated through the reanalysis of a number published imaging studies.

Keywords: multivariate analysis; representational analysis; inference; Bayesian modelling; fMRI

Introduction

How much do two brain activity patterns overlap? This question arises quite often in neuroimaging studies - for example when trying to establish to what degree two tasks engage overlapping or distinct neural processes. It is easy to establish whether two activity patterns correlate more with each other than chance. This can be simply accomplished by testing a set of correlation coefficients against zero. However, in the presence of substantial measurement noise, it is difficult to estimate the true size of the correlation coefficient (Fig. 1a).

The problem also occurs in multi-voxel pattern analysis (MVPA), where we may want to determine whether the differences between multiple items (i.e., different hand actions, arranged on a representation hyper-plane, Fig. 1b) are represented similarly across two conditions (e.g., observation vs. execution). By training a classifier on one condition and then testing it on the other, one can easily establish whether the representation in one condition holds any information about the other. However, the exact degree of the similarity between the representations (i.e., the angle between two representational hyper-planes) is difficult to estimate.

A typical frequentist approach (as taken in representational similarity analysis, RSA) is to derive corrected point-estimates for the correlation for each subject, and then use inter-subject variability (or bootstrap) to obtain variability measures for inference. We show here that these approaches start to fail quickly when signal-to-noise is low, as is often the case in fMRI. Instead we suggest a Bayesian approach to estimate the posterior distribution of the correlation coefficient from single-subject or group data, allowing for valid inference.

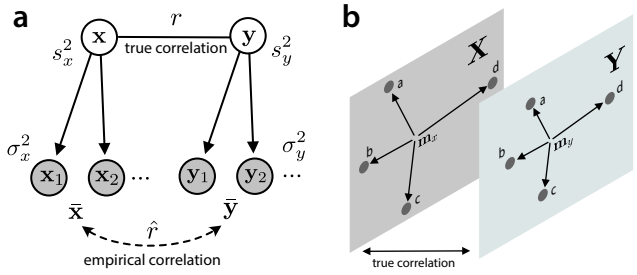

Figure 1: **a.** Basic case: determine the true correlation between two activity patterns from noisy measurements of each pattern. **b.** More complex case: determine the angle between two hyper-planes (gray planes) that distinguish between items (a-d) under two conditions ($\mathbf{X}, \mathbf{Y}$))

Methods and Results

Problem definition

We assume that the true activity patterns in the two condition $(\mathbf{x}, \mathbf{y})$ are normally distributed vectors with P voxels, and variance of s_x^2 and s_y^2. We have N independent measures $\mathcal{D} = \{\mathbf{x}_1, \mathbf{y}_1, ..., \mathbf{x}_N, \mathbf{y}_N\}$ each corrupted with $N(0, \sigma^2)$ noise. In general we are interested in the true correlation (or, when not subtracting the mean value across voxels, the true cosine similarity) between the two activity patterns:

$$r = \frac{\mathbf{x}^T\mathbf{y}/P}{\sqrt{s_x^2 s_y^2}}$$

A simple estimate can be obtained by plugging the estimated mean activity patterns $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ into Eq. 1. However, it has long been known (Spearman, 1987) that this estimate (Fig. 2a, gray dashed) underestimates the true correlation:

$$\hat{r}_{unc} = r\sqrt{\frac{s_x^2}{s_x^2 + \sigma_x^2}\frac{s_y^2}{s_y^2 + \sigma_y^2}} \quad (1)$$

Cross-block estimation. We can attempt to correct for this bias by using the measurements of $\mathbf{x}$ and $\mathbf{y}$ from different measurement blocks (which we assume are independent) to estimate the noise and signal variances and then correct the correlation estimate (Beaton et al., 1979). This procedure is similar to the procedure by which one can obtain unbiased distance estimates in RSA (Walther et al., 2016).

This estimator (Fig. 2a , blue line), however, has three undesirable features that make it unsuited for testing hypotheses of interest. (1) At low signal-to-noise levels (as typical in fMRI) the estimate of signal variance often becomes negative, forcing us to exclude many participants for inference. (2) The bias is not completely removed. (3) The estimates have very high variance, as it is not bounded between $[-1, 1]$.

Maximum-likelihood estimator. Using pattern-component modelling (PCM) (Diedrichsen, Yokoi, & Arbuckle, 2018), we can also derive an estimator of r that maximizes the data likelihood $p(\mathcal{D}|r, s_x^2, s_y^2, \sigma^2)$. This estimator is identical to the truncated (on [-1,1]) cross-block estimator. While much better behaved (Fig. 2a, red line), the estimator still shows substantial bias for typical noise levels of multivariate fMRI studies.

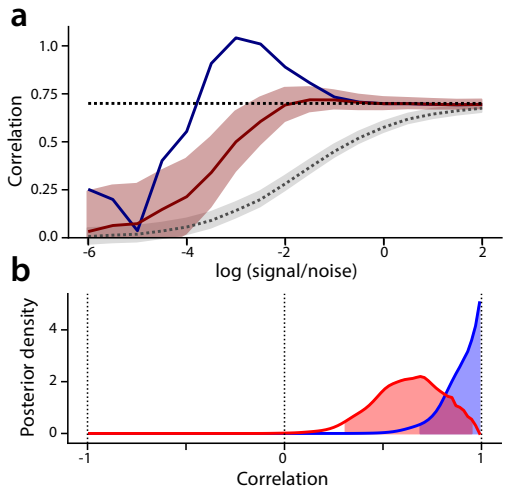

Figure 2: **a.** Mean correlation estimates for a true correlation of $r = 0.7$ depending on signal-to-noise ratio. Pearson correlation: gray dashed line; Cross-block estimator: blue line; Maximum-likelihood estimate: red line. Shaded area indicates SEM for 20 participants (not shown for cross-block estimate, as they are too big). **b.** Approximate posterior distribution for a single dataset with $r = 0.7$ (red) and $r = 1.0$ at a log signal-to-noise of -2 (blue). The 95% highest posterior density interval is indicated by the shaded area.

The Bayesian solution

Given the inadequacies of frequentist approaches to provide robust inference, we propose here to apply a Bayesian approach: Using a uniform prior on the correlation and Jeffrey's prior on the variance parameters ($p \propto \sigma^{-2}$), we can approximate the posterior distribution across all parameters using Markov-chain Monte-Carlo (MCMC) sampling. To enforce the [-1,1] bounds, sampling is performed in a fisher-z transformed correlation space, where the corresponding prior is $p \propto (1 - r^2)$. To enable group inference, MCMC is performed on a single model for the entire group of participants, using a common correlation parameter across participants, but separate signal and noise variance parameters.

One-sample problem. To test hypotheses about the size of correlation coefficients, we can approximate the marginal posterior using kernel-density estimation in fisher-z transformed space, which can then be transformed into the original space on the interval of $[-1, 1]$. This smoothed interval allows us to construct the 95% highest posterior density interval for the correlation coefficient (Fig. 2b).

In contrast to the central credibility interval (or an interval obtained by bootstrapping of the maximum-likelihood estimate), this confidence interval contains the correlation parameter of the true model with approximately the correct frequency across most noise and correlation levels.

Two-sample problem. The two-sample problem occurs when we compare correlations across two brain regions or two groups of subjects that have different signal-to-noise levels. We can approach this problem by sampling separate posterior distribution for both groups, and then compute the probability that $r_1 > r_2$ across any pairs of samples. Compared to the strongly biased inference using the maximum-likelihood estimate when the noise levels in the two samples differ, this approach leads to broadly correct inferences.

Discussion

An emerging principle in neural coding is that related processes share some overlap in terms of their neuronal states, but also have some differences in other dimensions. Such mixed encoding ensures that knowledge can be transferred from one condition to another while still allowing learning separate rules if necessary (Bernardi et al., 2020). The true angle between two representational hyper-planes (Fig. 1b) therefore becomes an important quantity to characterize the capacity of a region for common vs. separate encoding. Here we establish a robust Bayesian methods to estimate and obtain credibility interval for this statistics from data that contains substantial measurement noise. We demonstrate the feasibility of the approach through a re-analysis of published multi-variate fMRI studies (Ariani, Pruszynski, & Diedrichsen, 2022; Berlot, Prichard, O'Reilly, Ejaz, & Diedrichsen, 2018). The proposed models and inference procedure are implemented in a new version of the PCM toolbox (Diedrichsen et al., 2018).

Acknowledgments

Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC, RGPIN-2022-04692) to JD; BrainsCAN award to Western University.

References

Ariani, G., Pruszynski, J. A., & Diedrichsen, J. (2022, January). Motor planning brings human primary somatosensory cortex into action-specific preparatory states. *Elife*, *11*.

Beaton, G. H., Milner, J., Corey, P., McGuire, V., Cousins, M., Stewart, E., ... Little, J. A. (1979). Sources of variance of 24-hour dietary recall data: Implications for nutrition study designing and interpretation. *Am. J. Clin. Nutr.*, *32*(12), 2546–2559.

Berlot, E., Prichard, G., O'Reilly, J., Ejaz, N., & Diedrichsen, J. (2018). Ipsilateral finger representations in the sensorimotor cortex are driven by active movement processes, not passive sensory input. *J. Neurophysiol.*.

Bernardi, S., Benna, M. K., Rigotti, M., Munuera, J., Fusi, S., & Salzman, C. D. (2020, November). The geometry of abstraction in the hippocampus and prefrontal cortex. *Cell*, *183*(4), 954–967.e21.

Diedrichsen, J., Yokoi, A., & Arbuckle, S. A. (2018, October). Pattern component modeling: A flexible approach for understanding the representational structure of brain activity patterns. *Neuroimage*, *180*, 119–133.

Spearman, C. (1987). The proof and measurement of association between two things. by c. spearman, 1904. *Am. J. Psychol.*, *100*(3-4), 441–471.

Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., & Diedrichsen, J. (2016, August). Reliability of dissimilarity measures for multi-voxel pattern analysis. *Neuroimage*, *137*, 188–200.