Noise Ceiling Estimation for Representational Similarity Matrices

more about Quarto see https://quarto.org.

Representational Similarity Analysis (RSA) in a nutshell

The idea is to compare the representational geometry of some neural data with that of a model (Kriegeskorte, Mur, and Bandettini (2008)). For example, we stimulate participants with each several repetitions of different visual stimulus categories. In fMRI, we then run a GLM with separate regressors per k category, extract the voxel estimates and put them into a vector v_k of a particular ROI we are interested in. For MEEG, v_k is typically composed of the electrical potentials from the different channels (e.g., electrodes) of an grand average event-related potential (ERP) – an average over several dozens of repetitions or epochs, and the following is then estimated for each time point of the ERP separately (but see Section). We then correlate all v_k with each other to get a $k \times k$ (symmetric) representational similarity matrix (RSM) D (or pattern similarity matrix PSM) representing the neural data ($data\ RSM$). For simplicity, and for the sake of this work, this part of the analysis will be named *first stage* of an RSA.

However, the main interest lies usually on the results of a second stage of an RSA. The lower triangle of D is then (rank) correlated with a $k \times k$ *model RSM M*. Different M can be compared to each other, e.g. via the magnitude of their correlation with D. For group level inference, D can be averaged (but see Section) to $\bar{D}$, and the result can be (rank) correlated with M.

Excursion: Fisher-r-to-z transformation of correlation coefficients

Fisher r-to-z transformation (i.e., $z = arctanh(r)$ or Equation 1) for correlation coefficients (including those for noise ceilings) is often not reported in the RSA literature (exceptions, e.g.,

Carlin and Kriegeskorte ([2017])), and some packages even haven't implemented it for some reasons (e.g., see discussion here).

$$z = \frac{1}{2}ln\left[\frac{1+r}{1-r}\right] \tag{1}$$

For steps in which any r needs to be averaged (e.g., across samples, or across subjects), mean r is however biased, especially for $r > 0.5$. The numeric difference for low r can be treated as neglectable. However, due to the low costs of transformation, I will always implicitly assume that r values are transformed to z for each step that includes e.g., averaging or statistics. After that being done, for the sake of easier interpretability, it is then transformed back using $r = tanh(z)$ or Equation 2

$$r = \frac{e^{2z}-1}{e^{2z}+1} \tag{2}$$

Noise ceiling (NC) for (second stage) RSA

To get an estimate of how well the best possible model could perform, given the noise conditions in the data, noise ceilings (NC or ν) can be estimated (e.g., Lage-Castellanos et al. ([2019])). I found versions suitable for RSA. Both of them are applied to estimate NCs for second stage RSA.

Boundary procedure (group level inference)

The boundary procedure was for example used in Kaniuth and Hebart ([2022]) and Xie et al. ([2022]) and implemented in Nili et al. ([2014]). The lower ν_{lower} and upper bound ν_{upper} of the noise ceiling are estimated separately:

For the upper bound

1. Estimate an average D ($\bar{D}$) across participants.
2. Correlate the (lower triangular) of each participants D_i to $\bar{D}$, and retrieve $\nu_{u,i}$
3. Average across $\nu_{u,i}$, to get ν_{upper}

This is the upper bound, because it is an overestimated noise ceiling, as $\bar{D}$ and D_i share some information about one participant.

For the lower bound,

1. For each participant, estimate $\bar{D}^{(-i)}$ across participants, which is the average D but without participant i.

2. Correlate the (lower triangular) of each D_i to $\bar{D}^{(-i)}$, and retrieve $\nu_{l,i}$
3. Average across $\nu_{l,i}$, to get ν_{lower}

This is the lower bound, because $\bar{D}^{(-i)}$ does not have as much data as $\bar{D}$ (i.e., participant i). Typically, ν_l and ν_u bound span a region which is shown alongside with the outcome of the (second stage) RSA.

Split-half approach (single-participant or group level)

I have not found any practical example where split half noise ceiling ν_{sh} is used in the context of *RSA*, but rather for calculation of noise ceilings in other contexts. However, Lage-Castellanos et al. (2019) proposes this method also for RSA matrices, without going into detail. For typical (second stage) *RSA* (on the single participant level), the idea would roughly be the follows:

1. Split the trials per condition in half, and estimate from each respective half the corresponding *RSA*, $D_{i,1}$ and $D_{i,2}$
2. Correlate the (lower diagonal) of $D_{i,1}$ and $D_{i,2}$ (Equation 3).
3. Adjust this correlation coefficient for the reduced number of samples that enter the correlation in the split-half approach compared to a correlation that would use as many data as the full data set had for each partner that enters the correlation (Equation 4).

$$\nu_{sh,unadj} = \frac{cov(\hat{\beta}_1, \hat{\beta}_2)}{\sqrt{\sigma_{\hat{\beta}_1} \sigma_{\hat{\beta}_2}}} \tag{3}$$

where β are estimates from fMRI GLMs of the respective two splits.

$$\nu_{sh} = \sqrt{\frac{2\nu_{sh,unadj}}{\nu_{sh,unadj} + 1}} \tag{4}$$

However, see Section and Section , as it is unclear, which version is correct for Equation 4. Also note, that for group level, the adjustment factor might maybe not work, as the Spearman-Brown adjustment factor (Section) might be rather made for an individual level correlation value adjustment (TODO: check if this is correct). But maybe the method proposed in Section might work.

Proposed method: noise ceiling for participant-level RSMs (first stage RSA)

For the sake of clarity, we define any symmetric D_i for participant i as a matrix with its on-$(d_{q,q})$ and off-diagonal $(d_{q,p}^{(q \neq p)})$ elements as in Equation 5.

$$D_i = \begin{pmatrix} d_{1,1} & d_{1,2} & d_{1,3} & d_{1,4} \\ d_{2,1} & d_{2,2} & d_{2,3} & d_{2,4} \\ d_{3,1} & d_{3,2} & d_{3,3} & d_{3,4} \\ d_{4,1} & d_{4,2} & d_{4,3} & d_{4,4} \end{pmatrix} \tag{5}$$

We want to estimate a noise ceiling , that is, the maximal possible correlation value that can be achieved, given the noise conditions. In *RSMs* (first stage *RSA*) this is the maximal possible correlation value we can expect within the *RSM*. The on-diagonal elements offer themselves as natural noise ceiling, as their value of 1 is are the upper bound of what can be achieved, namely the correlation of the pattern vectors v_k with themselves. However, this value is overestimated, then even with pure noise in the data, the on-diagonals would be 1. However, we could use a split-half approach.

For instance, we have a D_i estimated on the full data and we want to estimate ν for one particular element $d_{q,p}^{(q \neq p)}$ of D_i, i.e., for one off-diagonal correlation coefficient. We split the data set in two halves, and estimate $D_{i,sh}$ by correlating $v_{k,1}$ (1st half) and $v_{k,2}$ (2nd half), which are e.g. the ERPs from the respective half of epochs, effectively making $D_{i,sh}$ asymmetric (Equation 6). (Similar as done here for a different purpose).

$$D_{i,sh} = \begin{pmatrix} d_{1,1}^{sh} & d_{1,2}^{sh} & d_{1,3}^{sh} & d_{1,4}^{sh} \\ d_{2,1}^{sh} & d_{2,2}^{sh} & d_{2,3}^{sh} & d_{2,4}^{sh} \\ d_{3,1}^{sh} & d_{3,2}^{sh} & d_{3,3}^{sh} & d_{3,4}^{sh} \\ d_{4,1}^{sh} & d_{4,2}^{sh} & d_{4,3}^{sh} & d_{4,4}^{sh} \end{pmatrix} \tag{6}$$

We need to make one assumption: the noise structure is the same in all k conditions.

Then we can treat the on-diagonals $d_{q,q}^{sh}$ as the noise ceiling of the off-diagonals $d_{q,p}^{(q \neq p),sh}$. The on-diagonals are the correlations within the same category, whereas the off-diagonals are correlations between categories. Our assumption is, that the correlation of one category with itself is the maximal possible correlation that can be achieved. Whereas $d_{q,q}$ were estimated to 1, as there was not another test set of v_k to be correlated with, $d_{q,q}^{sh}$ was estimated by two sets of v_k.

The interpretation of the noise ceiling is slighly different than the noise ceiling estimated for the correlation between D and M, which is typically for model comparison and states, how well a model can theoretically achieve. Here, it's how well a correlation coefficient can score between categories, and this becomes higher and higher, as more similar the categories are,

4

peaking in the highest possible correlation, when the categories the same (but the trials used for calculation are different).

Assuming equal noise structure through all categories, one could estimate $\nu_{sh,unadj}$ as average over all (r-to-z transformed and backtransformed) on-diagonals $d_{q,q}^{sh}$, i.e.:

$$\nu_{sh,unadj} = tanh\left(\frac{1}{k}\sum_{q}^{k} arctanh(d_{q,q}^{sh})\right) \tag{7}$$

The split-half approach could be repeated with e.g., 1000 different splits to get a more stable estimate of $\nu_{sh,unadj}$, i.e., stratified Monte Carlo splitting.

Now we need to adjust this value for the reduced amount of data in the split half approach (Section).

Adjustment coefficient

When calculating the split-half noise ceiling (ν_{sh}) via a correlation coefficient, the correlation estimate is biased downwards because the correlation does not have the same amount of data, then the original correlation had, for which the ν is to be estimated. If one has an abundance of data, and does not care about dropping half of it, one could just use the full data set for noise ceiling estimation, and then drop half of it. Then one has an accurate noise ceiling estimate for correlations for the D_i or D of the remaining trials with a potential M (i.e., the second stage and often main part of interest of the *RSA*). The same logic should count for ν_{sh} estimation for *PSM/RSM*s.

Spearman-Brown prophecy formula

The recent paper of Lage-Castellanos et al. (2019) mentions an adjustment for this underestimation (Equation 3 & Equation 4). They reference Huth et al. (2012) and Luking et al. (2017) for the ν_{sh} calculation. The origin of the adjustment factor (Equation 4) is not clear. Luking et al. (2017) name the *Spearman-Brown prophecy formula* to adjust for their split-half Pearson correlation coefficient used, writing the formula

$$\nu_{sh} = \frac{2\nu_{sh,unadj}}{1 + \nu_{sh,unadj}} \tag{8}$$

The origin of the square root in Equation 4 compared to Equation 8 is not clear. However, it inflates the noise ceiling estimation. Warrens (2016), to which the r package *splithalfr* refers to, and Pronk et al. (2022), also use Equation 8 and states, that it the method holds only

if σ_1^{sh} and σ_2^{sh} do not differ substantially, and if the lengths of the two halves do not differ significantly. Warrens (2016) also introduces alternative adjustment factors.

Maybe: is Equation 4 thinking it's an R^2 estimate instead of ρ, and therefore for some reasons use square root?

Huth et al. (2012) only briefly mentions noise ceiling estimation in the appendix, with reference to Hsu, Borst, and Theunissen (2004), which seems mathematical rather complex, and based on spike train data (also could not find an exact formula on which this is based on).

Wikipedia also describes the Spearman-Brown Formula as

$$\nu_{adj} = \frac{n\nu_{sh,unadj}}{1 + (n-1)\nu_{sh,unadj}} \tag{9}$$

where n is the factor by which the length of the data is changed, i.e., if reduced from 10 to 5, then it's 2. If $n = 2$ such as in split half noise ceiling estimation, then this reduced to Equation 8.

Empirical adjustment factor / projection factor

Here, an empirical adjustment factor is proposed. The idea is simple, and I don't find reasons why it should be less accurate (TODO: simulate?) than the analytic suggestion of Lage-Castellanos et al. (2019), of which the origin is not fully traceable right now.

We already estimated unadjusted split half noise ceilings in Section via Equation 7 as an average of on-diagonal correlation coefficients.

Now we could use the off-diagonal elements $d_{q,p}^{(q \neq p)}$ to estimate the adjustment factor. We know the original correlation values of the full-data RSM D_i. Further, we already estimated the split half data RSM $D_{i,sh}$ and its off-diagonal values $d_{q,p}^{(q \neq p),sh}$. We could now average (with r-to-z-transform and back) the off-diagonals of D_i and $D_{i,sh}$, respectively, and obtain r_f and r_p (the full and partial correlation coefficients) (Note that all are z values).

The adjustment factor β is then

$$\beta = \frac{r_f}{r_p}$$

We can simply multiply β with the $\nu_{sh,unadj}$ to get a ν_{sh} value, which can be used as estimate for noise ceiling.

$$\nu_{sh,adj} = \beta \times \nu_{sh,unadj}$$

Time-resolved or concatenated analysis?

WIP

The EEG literature is full of time-resolved second stage RSA (e.g., XXXREF, XXXREF, XXXREF). However, for first-stage RSA, one might be interested in the pattern similarity over a particular time window (e.g. post-stimulus onset) rather than each timepoint individually. Possible options, however not yet found in the literature are:

1. Concatenating electrode voltages over time and correlating concatenated vectors rather than only channel activations of one timepoint
2. Do the usual (first stage) RSA for each time point, and average (z-transformed) RSMs across time.

Sandbox

WIP

for RSA (2nd level) TODO: does it work even?

For instance, we have a *data RSM* and want to estimate the NC for the comparison to a *model RSM* (i.e., 2nd level of a typical *RSA*). One would split the data set into half, generate 2 *RSM*s, and correlate the off-diagonals of these *RSMs*. The resulting adjustment coefficient is the noise ceiling, but needs to be adjusted. Therefore, we calculate

References

Carlin, Johan D., and Nikolaus Kriegeskorte. 2017. "Adjudicating Between Face-Coding Models with Individual-Face fMRI Responses." *PLOS Computational Biology* 13 (7): e1005604. https://doi.org/10.1371/journal.pcbi.1005604.

Hsu, Anne, Alexander Borst, and Frédéric E. Theunissen. 2004. "Quantifying variability in neural responses and its application for the validation of model predictions." *Network (Bristol, England)* 15 (2): 91–109.

Huth, Alexander G., Shinji Nishimoto, An T. Vu, and Jack L. Gallant. 2012. "A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories Across the Human Brain." *Neuron* 76 (6): 1210–24. https://doi.org/10.1016/j.neuron.2012.10.014.

Kaniuth, Philipp, and Martin N. Hebart. 2022. "Feature-Reweighted Representational Similarity Analysis: A Method for Improving the Fit Between Computational Models, Brains, and Behavior." *NeuroImage* 257 (August): 119294. https://doi.org/10.1016/j.neuroimage.2022.119294.

Kriegeskorte, Nikolaus, Marieke Mur, and Peter Bandettini. 2008. "Representational Similarity Analysis – Connecting the Branches of Systems Neuroscience." *Frontiers in Systems Neuroscience.* https://doi.org/10.3389/neuro.06.004.2008.

Lage-Castellanos, Agustin, Giancarlo Valente, Elia Formisano, and Federico De Martino. 2019. "Methods for Computing the Maximum Performance of Computational Models of fMRI Responses." Edited by Jörn Diedrichsen. *PLOS Computational Biology* 15 (3): e1006397. https://doi.org/10.1371/journal.pcbi.1006397.

Luking, Katherine R., Brady D. Nelson, Zachary P. Infantolino, Colin L. Sauder, and Greg Hajcak. 2017. "Internal Consistency of Functional Magnetic Resonance Imaging and Electroencephalography Measures of Reward in Late Childhood and Early Adolescence." *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 2 (3): 289–97. https://doi.org/10.1016/j.bpsc.2016.12.004.

Nili, Hamed, Cai Wingfield, Alexander Walther, Li Su, William Marslen-Wilson, and Nikolaus Kriegeskorte. 2014. "A Toolbox for Representational Similarity Analysis." *PLOS Computational Biology* 10 (4): e1003553. https://doi.org/10.1371/journal.pcbi.1003553.

Pronk, Thomas, Dylan Molenaar, Reinout W. Wiers, and Jaap Murre. 2022. "Methods to Split Cognitive Task Data for Estimating Split-Half Reliability: A Comprehensive Review and Systematic Assessment." *Psychonomic Bulletin & Review* 29 (1): 44–54. https://doi.org/10.3758/s13423-021-01948-3.

Warrens, Matthijs J. 2016. "A Comparison of Reliability Coefficients for Psychometric Tests That Consist of Two Parts." *Advances in Data Analysis and Classification* 10 (1): 71–84. https://doi.org/10.1007/s11634-015-0198-6.

Xie, Siying, Stefanie Hoehl, Merle Moeskops, Ezgi Kayhan, Christian Kliesch, Bert Turtleton, Moritz Köster, and Radoslaw M. Cichy. 2022. "Visual Category Representations in the Infant Brain." *Current Biology* 32 (24): 5422–5432.e6. https://doi.org/10.1016/j.cub.2022.11.016.