Divergence modeling: Analyzing perceptual representations via stimulus similarity and information theory

Anthony D. Cate[1], Timothy J. Herron[1] & David L. Woods[1,2,3,4]
[1]Human Cognitive Neurophysiology Lab, VA Martinez, CA [2]Department of Neurology
[3]Center for Neuroscience and [4]Center for Mind and Brain University of California at Davis

GOALS

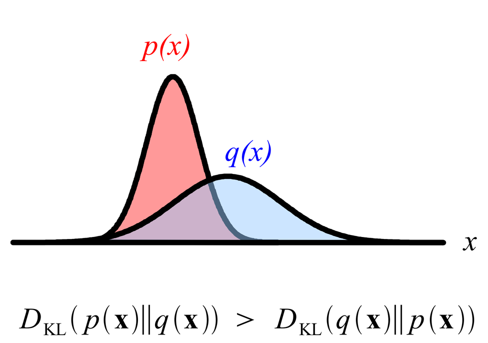
- Create a unified model of perceptual space from simple responses, as in MDS (multidimensional scaling)
- Accurately model stimulus tuning using Gaussian probability densities
- Characterize how cognitive processes alter representations, powerfully & simply

Why represent stimuli as normal probability density functions (PDFs)?

- Neurally plausible: can describe stimulus population codes (c.f. Deneve et al. 1999).
- Dissimilarity measure for PDFs can capture common non-Euclidean properties of behavioral and neural representations of stimulus dissimilarity: e.g. asymmetry and triangle inequality.

Natural measure of dissimilarity between PDFs: Kullback-Leibler divergence.

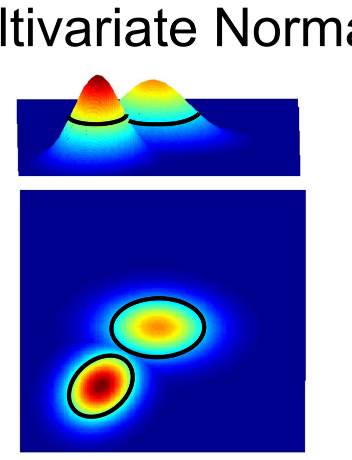
- Expresses amount of discrimination information, in sense of information theory.
- KL Divergence provides means modeling PDFs for entire perceptual space at once.



$D_{KL}(p(\mathbf{x})\|q(\mathbf{x})) > D_{KL}(q(\mathbf{x})\|p(\mathbf{x}))$

$$D_{KL}(p(\mathbf{x})\|q(\mathbf{x})) = E_{p(x)}\left[\ln\frac{p(x)}{q(x)}\right] = \int_X p(\mathbf{x})\ln\frac{p(\mathbf{x})}{q(\mathbf{x})}\,d\mathbf{x}$$

Multivariate Normal PDFs



$$\mu_1 = \begin{bmatrix}\mu_{1a}\\ \mu_{1b}\end{bmatrix} \qquad \delta = \mu_1 - \mu_2$$

For modeling, PDF means are **fixed** according to experimenter-defined values

$$\Sigma_1 = \begin{bmatrix}\sigma_{1a}^2 & \rho_{1ab}\sigma_{1a}\sigma_{1b}\\ \rho_{1ab}\sigma_{1a}\sigma_{1b} & \sigma_{1b}^2\end{bmatrix}$$

Covariance matrix values are **free parameters** for modeling

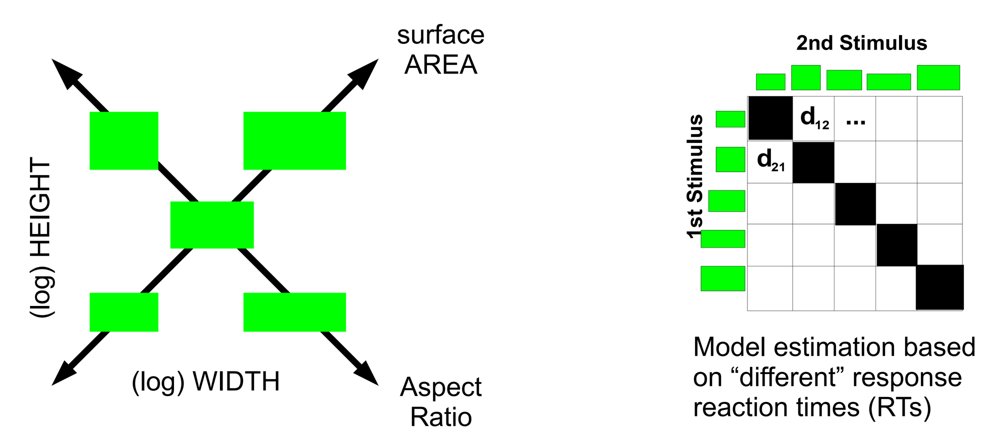
Kullback-Leibler Divergence

$$D_{KL}(\mathcal{N}_1\|\mathcal{N}_2) = \frac{1}{2}\left(\underbrace{\ln\frac{|\Sigma_2|}{|\Sigma_1|}}_{\Delta\text{ size}} + \underbrace{\mathrm{tr}\,\Sigma_1\left(\Sigma_2^{-1}-\Sigma_1^{-1}\right)}_{\Delta\text{ shape}} + \underbrace{\delta^T\Sigma_2^{-1}\delta}_{\Delta\text{ position}}\right)$$

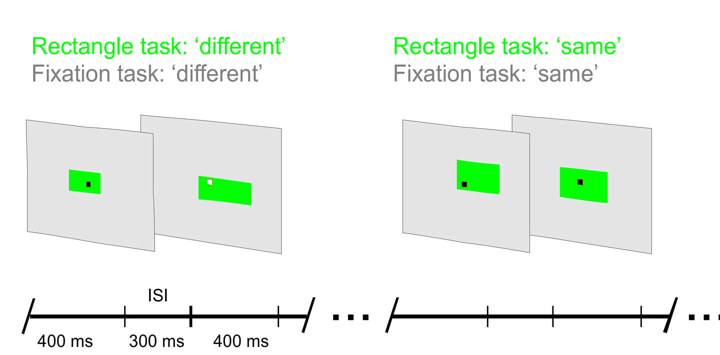
MODELING METHODS

Sample behavioral experiment:

- Effect of attention on rectangle perception
- Allowing asymmetric responses increases information available for modeling



Model estimation based on "different" response reaction times (RTs)



Max rectangle dimension = 5°

Screen position jittered on every presentation

In 2D stimulus space, need to model 3 PDF parameters:

σ_w, σ_h, and ρ_{wh}.

Important to have greater number of mean pairwise dissimilarities (20 here) than free parameters (15).

Reaction times (RTs) transformed into dissimilarity scores.

Attentive task: RTs logged, fit to normal cumulative density function, inverted about mean.

Unattentive (fixation) task: same, but not inverted. Greater dissimilarity in ignored rectangle pairs assumed to be more distracting, producing higher RTs.

Estimate stimulus PDF parameters

Minimize stress function by finding PDF parameters that let divergences match dissimilarities

Parameter fitting can be done using derivative-free (simplex) search, or with more precise gradient descent

Stress: $$\sigma_r = \sqrt{\frac{\sum_{i\neq j}\left(d_{ij} - D_{KL}(\mathcal{N}_i\|\mathcal{N}_j)\right)^2}{\sum_{i\neq j} d_{ij}^2}}$$

Two keys to optimization:

1) Initialize parameter search at good values

If assume that all covariance matrices are identical, D_{KL} becomes symmetric.

$$D_{KL}(\mathcal{N}_1\|\mathcal{N}_2;\Sigma) = \frac{1}{2}(\delta^T\Sigma^{-1}\delta) = D_{KL}(\mathcal{N}_2\|\mathcal{N}_1;\Sigma)$$

Can solve system of linear equations to find common covariance matrix that fits symmetrized dissimilarity measures.

$$d_{ij\,SYM} = \frac{1}{2}(d_{ij}+d_{ji})$$

2) Account for Riemannian curvature of parameter space during gradient descent

Here, $p(x)$ is the normal pdf with n parameters described by the vector $\boldsymbol{\theta}$

$$p(x;\boldsymbol{\theta}) \sim \mathcal{N}(\boldsymbol{\mu},\Sigma)$$

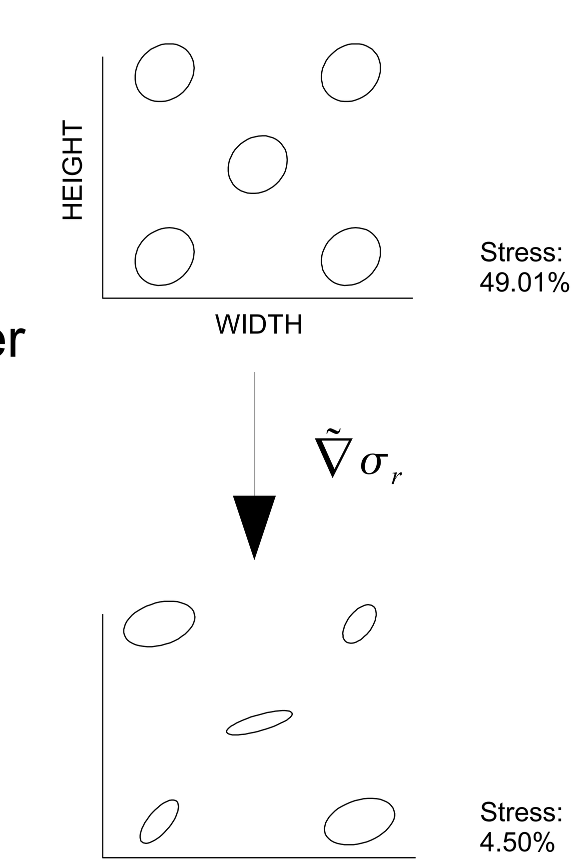
$$S = \{\boldsymbol{\theta} \in \mathbb{R}^n\}$$

$$G = \begin{bmatrix}\frac{\partial^2 \ln p(x)}{\partial^2\theta_1} & \cdots & \frac{\partial^2 \ln p(x)}{\partial\theta_1\theta_n}\\ \vdots & \ddots & \vdots\\ \frac{\partial^2 \ln p(x)}{\partial\theta_1\theta_n} & \cdots & \frac{\partial^2 \ln p(x)}{\partial^2\theta_n}\end{bmatrix}$$

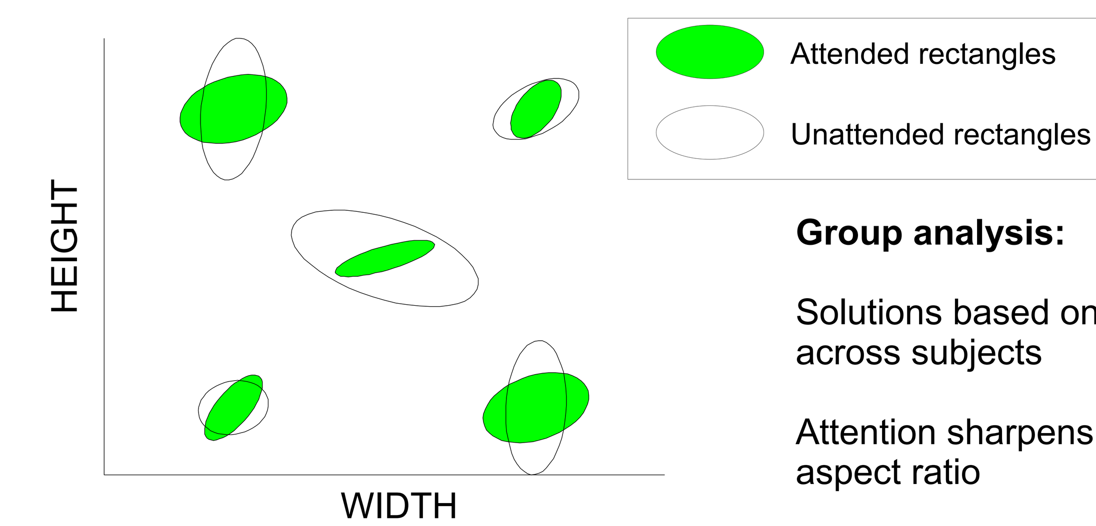
The parameter space S is not Euclidean; it is a Riemannian manifold whose curvature is described by the Fisher information (G)

$$\tilde{\nabla}\sigma_r = G^{-1}\nabla\sigma_r$$

Steepest descent follows the **natural (contravariant) gradient**



ANALYSIS METHODS



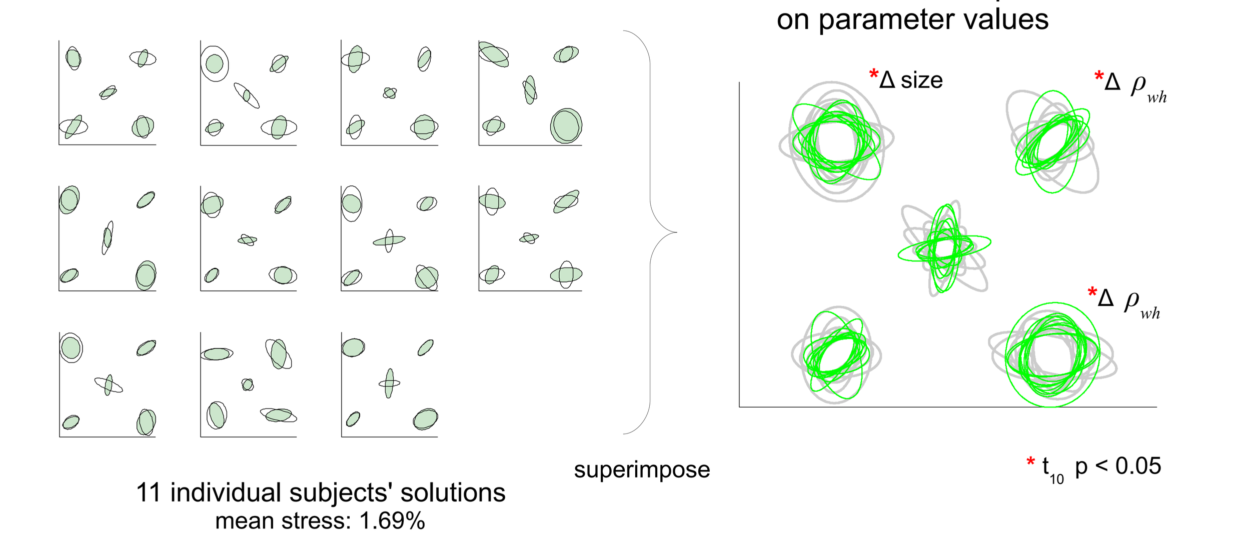
Group analysis:

Solutions based on mean data across subjects

Attention sharpens tuning to aspect ratio

Random effects analysis

Since data from different sessions transformed based on same normal distribution, can perform t- and F-tests on parameter values



11 individual subjects' solutions
mean stress: 1.69%

FUTURE WORK

- Apply to high-dimensional stimuli (e.g. faces)
- Estimate PDFs based on distribution of all RTs, not just means
- "Non-metric" modeling: preserve ordinality
- Incorporate decisional bias
- Interpolate results to estimate curvature of 2D stimulus parameter submanifold
- Analyze transformation that characterizes differences between manifolds for two experimental conditions