Updated Version of the Jern & Kemp (2013) Model

Working backwards, what is needed for simulation of exemplar generation is the probability of generating a stimulus y given exposure to members of the target category x_B:

$$p(y|x_B) =? \tag{1}$$

where x_B may be empty. Jern & Kemp's model achieves this using a generative process. Category members (x_A or x_B; more generally written as x_C) are assumed to have been generated using an underlying category distribution (specifically, multivariate normal):

$$x_C \sim Normal(\mu_C, \Sigma_C) \tag{2}$$

$p(y|x_C)$ is proportional to the candidate's density under the target category's distribution. Thus, obtaining the category distribution parameters (μ_C, Σ_C) is key for generation. This document describes how we compute these variables in a conjugate model.

Computing μ_C

Assuming (μ_C, Σ_C) are Normal-Inverse-Wishart distributed (unknown mean, unknown variance):

$$\mu_C = \frac{\kappa\mu_0 + n_C\bar{x_C}}{\kappa + n_C} \tag{3}$$

where:

- μ_0 is the prior mean along p dimensions. Here we set it to the middle of the space.
- κ is a scalar hyper-parameter, roughly weighting the importance of μ_0. κ must be greater than zero.
- n_C is the number of observations in x_C
- $\bar{x_C}$ is the sample mean along p dimensions

In the case of a populated class, μ_C ends up lying somewhere between μ_0 and $\bar{x_C}$, depending on κ_0 and n_C. In the case of an empty class, $n_C = 0$, Equation 3 reduces to $\mu_C = \mu_0$. Because we set μ_0 to the center of the space, this outcome is the same as if we had integrated over all possible μ_C. In practice, if $n_C = 0$, the model picks a stimulus at random from all candidates (uniform probabilities).

Computing Σ_D

Unlike μ_C, Σ_C cannot be computed considering only the members of category y. Instead, Σ_C is influenced both by the distribution of x_C and by members of other categories through Σ_D.

Σ_D is inferred based on the observed (empirical) category covariances C_y. We assume these covariances to be Wishart-distributed, and so Σ_D can be computed as:

$$\Sigma_D = \Sigma_0 + \sum_C C_C \tag{4}$$

Σ_0 is a p-by-p prior covariance matrix. We use a p-dimensional identity matrix I_p multiplied element-wise against a free parameter, ρ, controlling the amount of variance assumed by the prior:

$$\Sigma_0 = I_p \rho \tag{5}$$

Thus, categories are assumed to have some degree of variance along each feature (specified by ρ), but not are assumed to possess feature-feature correlations.

Computing Σ_C

Assuming (μ_C, Σ_C) are Normal-Inverse-Wishart distributed, Σ_C can be computed as:

$$\Sigma_C = [\Sigma_D \nu + C_C + \frac{\kappa n_C}{\kappa + n_C}(\bar{x_C} - \mu_C)(\bar{x_C} - \mu_C)^T](\nu + n_C)^{-1} \tag{6}$$

κ, $\bar{x_C}$, C_C, n_C, μ_0, are the same values as described above. ν is an additional free parameter, weighting the importance of Σ_D. ν must be greater than $p-1$. When x_b is empty, Equation 6 reduces to $\Sigma_C = \Sigma_D$.

Computing response probabilities $p(y|x_C)$

x are assumed to be drawn from a the distribution given by $Normal(\mu_C, \Sigma_C)$. Thus,

$$p(y) \propto Normal(y; \mu_C, \Sigma_C) \tag{7}$$

In practice, $p(y)$ is computed by first obtaining the relative density of every possible generation candidate y_i under the category distribution. The end probability is a normalization of these values:

$$p(x) = \frac{exp(\theta Normal(y; \mu_C, \Sigma_C))}{\sum_i exp(\theta Normal(y_i; \mu_C, \Sigma_C))} \tag{8}$$

where θ is a response determinism parameter.

Description of free parameters

- κ. Scalar, $\kappa > 0$. Weights the importance of μ_0 in inferring category μ_C.
- ρ. Scalar, $\rho > 0$. Sets the assumed variance in the domain prior, Σ_0.
- ν. Scalar, $\nu > p-1$. Weights the importance of Σ_D in inferring the domain Σ_C.
- θ. Scalar, $\theta > 0$. Response determinism parameter.