Goal

(italicized words identify optimization goals)

Wish to find W which has the *most* regular behavior over the *largest* contiguous subset of initial concentrations C_0, given gaussian disturbances of W. The contiguity may not be required, but seems to make sense (modelling a region which the cell can reliably keep the ambient concentrations within). We could take into account non-contiguous subsets, but that is not likely to be useful (it would be weird to have a motif which has similar behavior over different non-contiguous regions of ambient concentration).

Formulation

We define the norm over a matrix.

$$|W| : \mathbb{R}_{n\times n} \to \mathbb{R} = \sqrt{\sum_{i,j\in[1,n]} W_{ij}} \tag{1}$$

Given fixed parameters $W \in \mathbb{R}_{n\times n}, \sigma \in \mathbb{R}$, define a Gaussian distribution of weights.

$$D(W'|W,\sigma) = (2\sigma^2\pi)^{-2} e^{\frac{-(|W'-W|)^2}{2\sigma^2}} \tag{2}$$

This says "mutations modify the weights of the matrix W' a Euclidean distance from the original weight matrix W with probability according to the gaussian distribution with parameter σ." This may not be realistic, but it captures the idea that if a mutation occurs, it is likely to modify multiple weights at once if it modifies any at all.

We define the "behavior" of a network (C_0, W) as the concentration vector $C(T)$ at some given time T. This assumes linear differential equations modelling activation and inhibition rates with no external input or output.

$$C(t|C_0, W) = C_0 + \int_0^t W \cdot C(t)dt \tag{3}$$

This may not be appropriate; some networks may have *very* reliable periodic behavior, where the period simply changes slightly in response to disturbance, which would lead to vastly different $C(T)$. However, if we restrict ourselves to networks which reach some sort of equilibrium (which need not be defined or used at this time), we can use just $C(T)$. We define networks that reach equilibrium over some region $\mathcal{C}$ as networks which have similar $C(T)$ for some large T over $\mathcal{C}$. $T > 0$ and must be "large enough to allow the system to get to equilibrium." We can accomplish this by simply choosing a timescale.

Variability

Variability V of a network is defined using variation v with an initial concentration (C_0, W) over some stable $\mathcal{C}$ is defined. $\mathcal{C}$ is the aforementioned contiguous stable region of $C_0 \in \mathbb{R}^n$.

$$\mathcal{C}(W) \in \mathbb{R}^n \tag{4}$$

$$v(C_0, W', W) = \int_{C_0' \in \mathcal{C}(W)} |C_0' - C_0| \cdot |C(T|C_0', W') - C(T|C_0, W)| \tag{5}$$

$$V(C_0, W) = \int_{W' \in \mathcal{W}} D(W'|W, \sigma) \cdot v(C_0, W', W) \tag{6}$$

This is not what we want, even if it's the right idea – how do we find $\mathcal{C}(W)$ given W? Also, we want to make this a continuous measure over all C_0, not just those in $\mathcal{C}$, while maintaining the original objective. **This is what I'm currently stuck on.**

Finding stable networks X, then, is defined as an optimization problem. The most robust networks (the ones with the minimum variability) are (hopefully) likely to produce interesting network motifs.

$$X^* = \min_W \int_{C_0 \in \mathcal{C}(W)} V(C_0, W) \tag{7}$$

Simulation seems possible when the above issues are addressed.