Problem Set 3 (due Feb. 24)
Nonlinear networks

February 22, 2005

In this assignment we will study the dynamics of the N-neuron network

$$\frac{dx_i}{dt} + x_i = \left[b_i + \alpha x_i - \beta \sum_{j=1}^{N} x_j \right]^+ \tag{1}$$

Here α and β are positive and denote the strengths of self-excitation and global inhibition, respectively. The function $[u]^+ = \max\{u, 0\}$ is the rectification nonlinearity.

We will study this network in two different regimes. First, the amplification mode ($\alpha < 1$) and then the Winner-Take-All mode ($\alpha > 1$). Note that $\alpha = 1$ (the integrator mode) will be studied at a later time.

1. Amplification mode ($\alpha < 1$)

   (a) The importance of nonlinearity

   Analyzing a nonlinear network is more difficult than analyzing a linear network. It wouldn't be worth the trouble, unless the nonlinear network exhibits behaviors that are qualitatively different and interesting. Therefore it's important to make a comparison.

   Simulate the network described above, as well as its linear version (without the rectification) with 21 neurons numbered $i = -10..10$. For the inputs, use $b_i = 10 - |i|$. Explore the parameter space (α and β) and plot the steady state activation for both networks in the same plot (you can use the **subplot** MATLAB command).

   i. Find parameter sets for which 1, 3 and 7 neurons of the nonlinear network are active, respectively and submit the plots (for both networks) along with the values for α and β.

   ii. Comment on how α and β influence the selectivity (number of active neurons) of the two networks. Could the response of the nonlinear network be obtained by thresholding the linear network's response? Explain why or why not.
   We will now demonstrate this formally.

   iii. Show that the outputs of the rectified linear network (meaning the linear network whose output is only rectified after computing the steady state) at steady state can be written as

   $$x_i = \frac{1}{1 - \alpha} \left[b_i - \theta(N) \right]^+ . \tag{2}$$

   where

   $$\theta(N) = \frac{N\beta}{1 - \alpha + N\beta} \bar{b}_N \tag{3}$$

   and $\bar{b}_N$ is the mean of all N inputs. In other words, the inputs are thresholded at $\theta(N)$ and amplified by a factor of $1/(1 - \alpha)$. How does θ depend on α and β for large N?
   More intuitively, in the rectified linear model, the positive x_i are influenced by "inhibition" from the negative x_i (the effect is actually excitatory due to the double minus signs).

In contrast, in the nonlinear network, rectification is an integral part of the computation, not an afterthought.

We will show that the inactive neurons contribute no inhibition to the active neurons. It is as if the inactive neurons have been excised from the network.

(b) Global inhibition sets a dynamic threshold in the nonlinear network

Now we consider the nonlinear network again.

i. Suppose that the inputs to the nonlinear network are ordered as $b_1 \geq b_2 \geq \ldots \geq b_N$, and at least one is positive. Assume that k neurons are active and show that there is a unique steady state

$$x_i = \frac{1}{1 - \alpha}[b_i - \theta(k)]^+ \tag{4}$$

where

$$\theta(k) = \beta \sum_{j=1}^{k} x_j = \frac{\beta}{1 - \alpha + k\beta} \sum_{j=1}^{k} b_j \tag{5}$$

Hint: consider the cases $x_i = 0$ and $x_i > 0$ separately.

This means that the gain of the common mode is $1/(1 - \alpha + k\beta)$. Note that it depends on the number of active neurons. This illustrates a critical point, which is that a network with threshold nonlinearity cannot be characterized by a unique gain, unlike a linear system. Instead, the gain depends on the active set.

ii. Use this result to show that k must satisfy

$$b_k > \frac{\beta}{1 - \alpha + \beta k} \sum_{i=1}^{k} b_i > b_{k+1} \tag{6}$$

iii. Prove that k exists and is unique.

iv. What is the ratio $(x_i - x_j)/(b_i - b_j)$, if neurons i and j are both active? This ratio is the differential gain. Show that this is bigger than the gain of the common mode.

v. Show that only the top neuron is active in the limit as $\alpha \to 1$.

2. Lateral inhibition mediates competition: Winner-Take-All ($\alpha > 1$)

(a) First, we will perform a local stability analysis. In the neighborhood of a fixed point with k active neurons, we can linearize the dynamics of the active neurons to obtain

$$\frac{dx_i}{dt} + x_i = b_i + \alpha x_i - \beta \sum_{j=1}^{k} x_j \tag{7}$$

Alternatively, we can write the matrix-vector form

$$\frac{dx}{dt} + x = b + Wx \tag{8}$$

where $W = \alpha I - \beta \mathbf{1} \mathbf{1}^T$, I is the identity matrix with ones on the diagonal, and $\mathbf{1}$ is the column vector consisting of all ones, and the outer product $\mathbf{1} \mathbf{1}^T$ is the matrix consisting of all ones. Find the k eigenvalues and eigenvectors of W.

Hint: One eigenvector is the common mode $\mathbf{1}$ and the rest are differential modes (vectors whose components sum to zero).

Are the differential modes stable or unstable for $\alpha > 1$? Derive the values of β, and k for which the common mode is stable.

(b) You should have shown that all differential modes are unstable in this case for $\alpha > 1$. Consequently any stable fixed point must have only a single neuron active (winner-take-all).

    i. Show that the activity of this single neuron is given by

$$x_i = \frac{b_i}{1 - \alpha + \beta} \tag{9}$$

    What happens if $\alpha > 1 + \beta$?

    ii. Show that this is a self-consistent solution for any i from 1 to k, where

$$b_k \geq \frac{(1 - \alpha + \beta)b_1}{\beta} > b_{k+1} \tag{10}$$

    In other words, any neuron whose input is greater than $(1 - \alpha + \beta)/\beta$ times the largest input can end up the winner, depending on the initial conditions of the dynamics. This phenomenon of the existence of multiple stable fixed points is called multi-stability for short.

    iii. Hysteresis.

    Consider a 2-neurons winner-take-all network with $b_1 = 1$. Determine the separation d of the two branches of the hysteresis-curve of this network as a function of α and β. Note: the hysteresis curve is given by plotting the winning neuron as a function of b_2. Imagine that b_2 is alternatively increased and decreased so as to induce switching.