MIT Department of Brain and Cognitive Sciences
9.641J, Spring 2005 - Introduction to Neural Networks
Instructor: Professor Sebastian Seung

Problem Set 4 (due March 3)
Lyapunov functions

March 1, 2005

In lecture, you were told that the stability of

$$\dot{x}_i + x_i = \left[b_i + \sum_j W_{ij} x_j \right]^+$$

could be analyzed in the nonnegative orthant using the Lyapunov function

$$L(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T(\mathbf{I} - \mathbf{W})\mathbf{x} - \mathbf{b}^T\mathbf{x}.$$

In this problem you will prove that L is a Lyapunov function for the specific case of the "winner-take-all" network

$$\dot{x}_i + x_i = \left[b_i + \alpha x_i - \beta \sum_j x_j \right]^+ .$$

1. Lyapunov function for the WTA network.

   (a) Specialize the above general expression for L to the WTA dynamics.

   (b) Prove that L is nonincreasing $(dL/dt \leq 0)$ on trajectories of the dynamics, with equality only at steady states of the dynamics.
   *H*int: $\frac{dL}{dt} = \sum_i \frac{\partial L}{\partial x_i}\dot{x}_i$

   (c) Prove that L is lower bounded if $\alpha < 1 + \beta$ in the nonnegative orthant, and is not lower bounded if $\alpha > 1 + \beta$.

   (d) Prove that L is *radially unbounded* $(L(\mathbf{cx}) \to \infty$ as $c \to \infty)$ for $\alpha < 1 + \beta$ and x in the nonnegative orthant. This completes the proof that L is a Lyapunov function of the network dynamics.

   *N*ote: In class we talked about co-positivity of the $I - W$ matrix which is a sufficient condition for L to be radially unbouded.

2. 2-neurons network

   Remember what you have learned in class for the general case of N neurons. This network architecture has three distinct regimes: (a) Weak excitation ($\alpha < 1$) could lead to k active neurons depending on how well-separated the inputs are. (b) Strong excitation ($\alpha > 1$) can lead to a winner-take-all operation and a single active neuron for well-separated inputs. (c) There exists a third case which is called the integration regime for $\alpha = 1$ for which only the maximally activated neuron is active.

   (a) Specialize the general expression for L to the WTA dynamics with two neurons. You will have to refer to this equation later.

(b) Make an input phase-diagram for each network regime, *i.e.* for each of the conditions $\alpha < 1$, $\alpha = 1$, $1 < \alpha < 1+\beta$, $\alpha > 1+\beta$, make a plot of the input space (b_1 vs b_2) and characterize the different regions for which x_1 and/or x_2 can be active.

(c) Study the Lyapunov function associated with each of those regimes and show how they relate to the phase-diagram you drew previously. Explore 'representative' values of α and β and in each case, describe the shape of L, explain how many possible steady-states there are and if those are stable or not. Show the effect of changing the network inputs on L. We expect you to submit plots containing the Lyapunov function L and the trajectories followed by the network dynamics (you can use the 'mesh' command to plot L and the 'hold' command to superimpose the two plots). You should also submit the parameters you used to simulate your particular dynamics.

(d) You showed in problem 1 that L is not a Lyapunov function of the dynamics if $\alpha > 1+\beta$. What happens to the network in this case? Justify your answer using both a plot as well as mathematical arguments. .

3. Unconditional MAX behavior

In the following you will show that when the network activities are all initialized to 0 (*i.e.* $x(0) = 0$), the network **always** selects the unit that receives the maximum input as the winner for $\alpha > 1$.

(a) Show that by changing variables, the network equation can be written as:

$$\dot{u_i} + u_i = b_i + \sum_j W_{ij}[u_j]^+$$

(b) Show that if $u_i = u_j$ then $\frac{d}{dt}(u_i - u_j) = b_i - b_j$.

(c) Show that the only possible 'winner' has to be the one receiving the maximal input.