(e1) (7 pts) Compute the expectation of the complete log-likelihood $p(\boldsymbol{w}, \boldsymbol{y}; \boldsymbol{\mu}, \lambda, \beta)$ based on $p(\boldsymbol{w}|\boldsymbol{y}; \boldsymbol{\mu}_{old}, \lambda_{old}, \beta_{old})$. Note that you can simplify the notation using $\boldsymbol{u}_N^{(old)}, \boldsymbol{\Lambda}_N^{(old)}$.

(e2) (8 pts) Maximize $\boldsymbol{\mu}, \lambda, \beta$ in (e1) to obtain $\boldsymbol{\mu}_{new}, \lambda_{new}, \beta_{new}$.

(f) Approximate the posterior $p(\boldsymbol{w}|\boldsymbol{y})$ with the variational mean-field inference $p(\boldsymbol{w}|\boldsymbol{y}) \approx \prod_{i=1}^{M} q(w_i|\boldsymbol{y})$.

(f1) (7 pts) Find the functional form for $q(w_i|\boldsymbol{y})$.

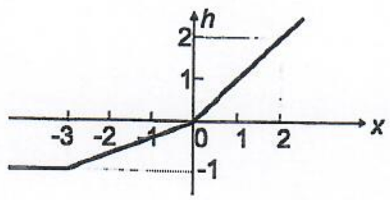
(f2) (8 pts) Provide fixed-point update equations for their parameters.

(g) (5 pts) Provide a procedure to perform combined variational inference and learning using the results in (f2) and (e1). Define your notation clearly.

[Note: I would encourage you to run computer simulations after the exam to see how this compares against the results obtained by the EM]

2. (8 pts) Enumerate and describe the activation functions, as many as you know.

3. (6 pts) Use the **maxout** unit to design the activation function below.



4. (8 pts) Enumerate and describe the pooling functions, as many as you know.

5. (8 pts) What are the major reasons that contribute to the success of convolutional neural networks?

6. (21 pts) For the following algorithm,

Algorithm Policy Iteration.

Initialize $V^{(0)}$ arbitrarily.

for $n = 1, 2, \ldots$ until termination condition do.

$\pi^{(n+1)} = \mathcal{G}V^{(n)}$

$V^{(n+1)} = \left(\mathcal{T}^{\pi^{(n+1)}}\right)^k V^{(n)}$, for integer $k \geq 1$.

end.

where $\mathcal{T}^{\pi}$ is defined as $[\mathcal{T}^{\pi}V](s) = \mathbb{E}_{s'|s,a=\pi(s)}[r + \gamma V(s')]$, and $\mathcal{G}$ is a greedy mapping function.

(a) (6 pts) Show that the value $V^{(n+1)}$ will converge when k is sufficiently large, assuming $\gamma < 1$. (Hint: policy evaluation)

(b) (6 pts) Show the convergence of $V^{(n+1)}$ assuming $\gamma < 1$, when n is sufficiently large, regardless of k. (Hint: value iteration upon $k = 1$, and policy iteration upon $k = \infty$)