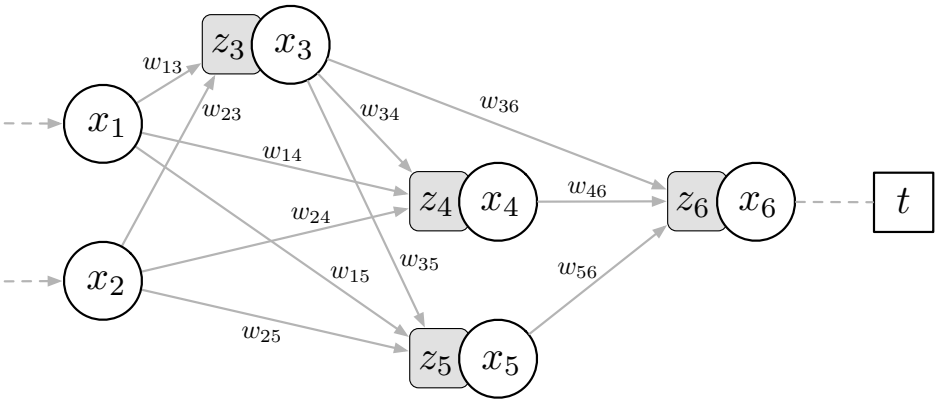
Gradient Descent and Error Backpropagation

Question 1: Backpropagation by Hand

This question involves the neural network shown below, with two input neurons, three hidden neurons, and one output neuron. For neuron j, its input current is z_j, its bias is b_j, and its activity is denoted x_j. All neurons use the same activation function, $\sigma(\cdot)$, except for the output neuron, which uses $\rho(\cdot)$. The connection weight **from** neuron i **to** neuron j is denoted w_{ij}. For example, the activity of neuron 5 can be calculated using, $x_5 = \sigma\left(w_{15}x_1 + w_{25}x_2 + w_{35}x_3 + b_5\right)$.



a) For the network above, suppose your cost function is the "Itakura-Saito distance", defined as

$$L(x,t) = \frac{t}{x} - \ln\frac{t}{x} - 1\,,$$

and the activation function for the output node is $\rho(z) = z^2$. Derive an expression for $\frac{\partial L}{\partial z_6}$, the gradient of the cost with respect to the input current to the output neuron. Show your work, and simplify your answer as much as possible.

b) Write down an expression for the update you would apply to the connection weight w_{46}. You may use your answer to question a. Be as explicit as possible.

c) Write down an expression for the update you would apply to the bias b_6. You may use your answer to question a. Be as explicit as possible.

d) Write an expression for $\frac{\partial L}{\partial z_3}$. Again, you may use your answer to question a. Show your work and simplify the formula as much as possible.

Submit your answers as a PDF file in Kritik. Make sure that nothing in the file can be used to identify you. Remember, the peer-assessment process is anonymous.