# CS5340 Lab 2 Part 2: Parameter Learning


```
Name: Niharika Shrivastava
Email: e0954756@u.nus.edu
Student ID: A0254355A
```


1. `_learn_node_parameter_w()`

- Taking a derivative of L wrt to all weights and equating to 0 gives us I+1 equations. We can write this in the form of matrix multiplications to solve it efficiently. Derivation provided in the end.
- Linear equation to be solved is:

```
Ax = B, where

A: Coefficient matrix of (I+1)*(I+1) -> observation of I parents with 1 bias in a square matrix
B: Output matrix of 1 * (I+1) -> observation of node multiplied by observation of each parent
x: Output weight parameters of 1 * (I+1)
```

2. `_learn_node_parameter_var()`

- Derivation provided in the end.

3. `_get_learned_parameters()`

- Construct a DGM using given nodes and edges.
- For each node, find its parents observations along with the node's observations and learn the weights.
- Learn the variance given the weights and all observations.
- Construct the required dictionary. `weight[0]` --> bias

NOTES

$$\frac{\partial L}{\partial w_{u0}} = \sum_{n=1}^{N} x_{u,n} - \sum_{n=1}^{N} (w_{u1} x_{u1,n} + \dots + w_{uc} x_{uc,n} + w_{u0}) = 0$$

$$\frac{\partial L}{\partial w_{u1}} = \sum_{n=1}^{N} x_{u,n} \cdot x_{u1,n} - \sum_{n=1}^{N} (w_{u1} x_{u1,n} \cdot x_{u1,n} + \dots + w_{uc} \cdot x_{uc,n} \cdot x_{u1,n} + w_{u0} \cdot x_{u1,n}) = 0$$

$$\vdots$$

$$\frac{\partial L}{\partial w_{uc}} = \underbrace{\sum_{n=1}^{N} x_{u,n} \cdot x_{uc,n}}_{B} - \underbrace{\sum_{n=1}^{N} (w_{u1} \cdot x_{u1,n} + \dots + w_{uc} \cdot x_{uc,n}) \cdot x_{uc,n}}_{Ax} = 0$$

$\rightarrow (C+1)$ equations. $\Rightarrow$ Solve $\boxed{Ax = B}$ using matrices.

$$\underbrace{\begin{bmatrix} \sum_{n=1}^{N} x_{u,n} \\ \sum_{n=1}^{N} x_{u,n} \cdot x_{u1,n} \\ \vdots \\ \sum_{n=1}^{N} x_{u,n} \cdot x_{uc,n} \end{bmatrix}}_{B}{}_{(C+1)(1)} = \underbrace{\begin{bmatrix} \sum_{n=1}^{N} (1) & \sum_{n=1}^{N} x_{u1,n} & \dots & \sum_{n=1}^{N} x_{uc,n} \\ \sum_{n=1}^{N} x_{u1,n} & \sum_{n=1}^{N} x_{u1,n} \cdot x_{u1,n} & \dots & \sum_{n=1}^{N} x_{uc,n} \cdot x_{u1,n} \\ \vdots & & & \\ \sum_{n=1}^{N} x_{uc,n} & \sum_{n=1}^{N} x_{u1,n} \cdot x_{uc,n} & \dots & \sum_{n=1}^{N} x_{uc,n} \cdot x_{uc,n} \end{bmatrix}_{(C+1)\cdot(C+1)}}_{A} \underbrace{\begin{bmatrix} w_{u0} \\ w_{u1} \\ w_{u2} \\ \vdots \\ w_{uc} \end{bmatrix}_{(C+1)(1)}}_{x}$$

Let $Y = \begin{bmatrix} 1 & x_{u1,1} & x_{u2,1} & \dots & x_{uc,1} \\ 1 & x_{u1,2} & x_{u2,2} & \dots & x_{uc,2} \\ \vdots & & & & \\ 1 & x_{u1,N} & x_{u2,N} & \dots & x_{uc,N} \end{bmatrix}_{N \times (C+1)}$ $\Rightarrow$ $B = (\text{outputs}) \times Y$, $A = Y^T \cdot Y$

$$L = \sum_{n=1}^{N} \left\{ -\tfrac{1}{2} \log(2\pi\sigma_u^2) - \frac{1}{2\sigma_u^2} \underbrace{\left(x_{u,n} - \left(\sum_{c \in x_{\pi_u}} w_{uc} x_{uc,n} + w_{u0}\right)\right)^2}_{a} \right\}$$

Let $\sigma_u^2 = x$

$$\Rightarrow L = -\frac{1}{2} \log(2\pi x) \cdot N - \frac{1}{2x} \sum_{n=1}^{N} a \Rightarrow \frac{\partial L}{\partial x} = -\frac{1}{\cancel{2}} \times \frac{N}{\cancel{2\pi} x} \cdot \cancel{2\pi} + \frac{\sum_{n=1}^{N} a}{\cancel{2} x^{2}} = 0$$

$$\Rightarrow \frac{\sum_{n=1}^{N} a}{x} = N \Rightarrow \boxed{x = \sigma^2 = \left(\frac{\sum_{n=1}^{N} a}{N}\right)}$$