1 Question 1

This is my answer to question 1. Below is an equation:

$$W = \underset{W \in \mathcal{W}}{\arg\min} \mathcal{L}(E_W(X, Y), \mathcal{S}) \tag{1}$$

2 Question 2

This is my answer to question 2, with a reference to Fig. 1

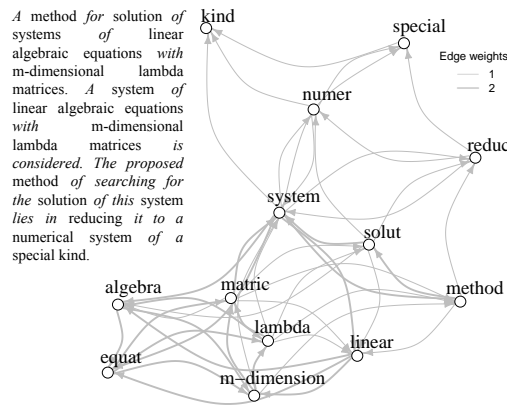

Figure 1: This is a caption.

3 Question 3

This is my answer to question 3, with a citation [1].

References

[1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008, 2017.