<1> for RANSAC algorithm,

number of iterations $= \frac{\log(1-p)}{\log(1-w^m)}$

where p is probability of success.

w^m is probability of choosing a subset with no outliers.

$\therefore$ number of itr $= \frac{\log(1-0.95)}{\log(1-0.5^4)}$

$= \frac{\log(0.05)}{\log(0.9375)} = 46.4177$

Since, no of iterations cannot be fractional,

$\therefore$ no of iterations $= 47$

<2> $h^1 = \sigma(w^1 x)$

$h^2 = \sigma(w^2 h^1)$

$f(x) = \langle w^3, h^2 \rangle$

$$\frac{\partial f}{\partial w^1_{i,j}} = \frac{\partial f}{\partial h^2} \cdot \frac{\partial h^2}{\partial h^1} \cdot \frac{\partial h^1}{\partial w^1_{ij}}$$

$$= w^3 (h^2(1-h^2)w^2)\, h^1(1-h^1)\, x_j$$

<3> $\Delta_{ij}^{(2)} := \Delta_{ij}^{(2)} + \delta_i^{(3)} * (a^{(2)})_j$

in ~~and~~ vector form;

$\Delta^{(2)} := \Delta^{(2)} - \alpha * \delta_i^{(2)} * a^{(2)}$

~~where~~

<4) No inputs = d.

No of hidden units = M.

No of hidden layers = 1.

No of output units = c

Total number of weights = $d \times M + M \times c$

$= dM + Mc = M(d+c)$

Total number of bias = ~~d~~ $M + c$

∴ Total number of weights & bias.

$M(d+c) + (M+c)$

$= M(d+c+c) = M(d+2c)$

$\downarrow$

expression for total weights and bias.

<6> <a> The problems caused by scale symmetry are:-

→ ill conditioning
→ saddle points
→ multiple local minima

⇒ saddle points:- a saddle point is a point where the gradient is same for all sides A neural network with symmetry will not learn any thing and will not use previous info to move in correct directions. Hence, it may take a really long time to converge or may not converge at all.

⇒ Multiple local minima - if there are multiple local minima, the network will find ~~se~~ · one local minima and then get stuck there as it will not learn any thing to move it to global minima.

<b> Another type of weight space symmetry is permutation symmetry. This gives rise to multiple equivalent global minima. of loss function. and first order saddle. points located in path between global minima. This causes the algorithm to independently learn same pattern multiple times, leading to increase in amount. of time and data required for training. This also creates opportunities for biases to creep in.