A1) Loss = L(y, label)
Skip Connection Weight = W_{skip}
Output vector = y
Input vector = x
hidden layer output = k

Final Output ⇒ $y = \cancel{f(k + W_{skip} x)}\ f(k + W_{skip} x)$

f is activation function.

$$\frac{\partial L}{\partial W_{skip}} = \frac{\partial L}{\partial y} * \frac{\partial y}{\partial W_{skip}}$$

$$\cancel{= f'(k + W_{skip} x) * x^T}$$

$$= \frac{\partial L}{\partial y} * f'(k + W_{skip} x) * x^T$$

A2) For 1st derivative term → W independent elements.

H is symmetric → W (diagonal elements) + $\frac{W(W-1)}{2}$ (off diagonal element).

Combinging these $W + W + \frac{W(W-1)}{2}$

$$= \frac{W(W+3)}{2}$$

A3) Convex hull of $x_i = \sum \alpha_i x_i$, $\sum \alpha_i = 1$
Convex hull of $z_i = \sum \beta_i z_i$, $\sum \beta_i = 1$

$f_1(x) = w^T(\sum \alpha_i x_i) + w_0 = \sum(\alpha_i(w^T x_i + w_0))$

$f_2(x) = w^T(\sum \beta_i z_i) + w_0 = \sum(\beta_i(w^T z_i + w_0))$

If convex hull intersect then $\exists\, \alpha_i\, \&\, \beta_i$

s.t. $f_1(x) = f_2(x)$

(i) Suppose ~~the~~ their convex hull intersect & they are linearly separable.

Hence $\sum(\alpha_i(w^T x_i + w_0)) = \sum(\beta_i(w^T z_i + w_0))$ — ①

but $w^T x_i + w_0 > 0$ & $w^T z_i + w_0 < 0$ as given so ① cannot be true hence contradiction hence they are not separable.

(ii) As they are linearly separable.

$w^T x_i + w_0 > 0$ & $w^T z_i + w_0 < 0$

~~hence $\sum w^T(\beta_i z_i + w_0) \neq$~~

hence $\sum(\alpha_i(w^T x_i + w_0)) \neq \sum(\beta_i(w^T z_i + w_0))$

~~whenever~~ whenever α_i & β_i are non trivial.

$\Rightarrow f_1(x) \neq f_2(x)$ hence the 2 sets are disjoint