Exercise 1.

a) $\max_{x} L = w^T x + b - \lambda x^T x$

$$\frac{\partial L}{\partial x} = w - 2\lambda x \overset{!}{=} 0 \Rightarrow \boxed{x^* = \frac{w}{2\lambda}}$$

b) $\max_{x} L = \log P(w_c | x) + \log P(x)$

$$= w^T x + b + \log \frac{1}{\sqrt{\det(2\pi\Sigma)}} + -\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)$$

$$\frac{\partial L}{\partial x} = w - \Sigma^{-1}(x-\mu) \overset{!}{=} 0.$$

$$\boxed{x^* = \Sigma w + \mu}$$

c) $\max_{z \in Z} L = \underbrace{\log P(w_c | g(z))}_{= w^T \underset{Az + c}{x} + b} - \lambda \|z\|^2$

$$L = w^T A z + w^T c + b - \lambda z^T z$$

$$\frac{\partial L}{\partial z} = A^T w - 2\lambda z \overset{!}{=} 0 \Rightarrow \boxed{z^* = \frac{A^T w}{2\lambda}}$$

$$\boxed{x^* = A z^* + c = \frac{A A^T w}{2\lambda} + c}$$

d) if $AA^T = I$, $c = \vec{0}$,

the ~~protoy~~ prototype x^* from regularizers and generator are the same.

when $\Sigma w + \mu = \frac{w}{2\lambda}$

the prototype x^* from regularizers and experts are the same.

Exercise 2

Exercise 2.

~~because~~ $f(x) = w^T x = \sum_i w_i x_i$

a) sensitivity analysis:

$$\frac{\partial f}{\partial x_i} = w_i \Rightarrow R_i(x) = \left(\frac{\partial f}{\partial x_i}\right)^2 = w_i^2 \Rightarrow R = \begin{bmatrix} w_1^2 \\ w_2^2 \\ \vdots \\ w_d^2 \end{bmatrix}$$

$$\sum R_i(x) = \sum_i w_i^2$$

b) Taylor decomposition at root point $\tilde{x} = 0$

because $\tilde{x} = 0 \Rightarrow \tilde{x}_i = 0$

$$R_i(x) = \underbrace{\frac{\partial f}{\partial x_i}\Big|_{x=\tilde{x}}}_{=w_i} \cdot \underbrace{(x_i - \tilde{x}_i)}_{=x_i}$$

$$= w_i x_i$$

or because $f(tx) = w^T t x = t w^T x = t f(x)$

$f(0) = 0$.

choose $\tilde{x} = \lim_{t \to 0} t x$

$$R_i(x) = \frac{\partial f}{\partial x_i}\Big|_{\tilde{x}} \cdot x_i = w_i x_i$$

c) because $f(tx) = w^T t x = t w^T x = t f(x) \Rightarrow f(0) = 0$

and x have not limitation on its domain of definition

so the root point should be $x = 0$

which included in x's domain of definition.

$$\Rightarrow R_i = \frac{\partial f}{\partial x_i}\Big|_{x=\tilde{x}} (x_i - \tilde{x}_i)$$

$$= w_i x_i$$

Exercise 3.

a) because we have LRP - $\alpha_1\beta_0$ have:

$$R_j = \sum_k \frac{a_j w_{jk}^+}{\sum a_j w_{jk}^+} R_k.$$

in network (a):

$$R_2 = \frac{a_3 \cdot 1}{\sum a_3 \cdot 1} y = y$$

$$R_1 = \frac{a_1 \cdot 1}{\sum a_1 \cdot 1} R_2 = y$$

in network (b):

$$R_2 = \frac{a_4 \cdot 1}{\sum a_4 \cdot 1} y = y$$

$$R_1 = \frac{a_2 \cdot 1}{\sum a_2 \cdot 1} y = y.$$

in network (c)

$$R_2 = \frac{a_3 \cdot 0.5}{a_3 \cdot 0.5} y = y$$

$$R_{11} = \frac{a_1 \cdot 1}{a_1 \cdot 1 + a_2 \cdot 1} R_2 = \frac{a_1}{a_1 + a_2} y$$

$$R_{12} = \frac{a_2 \cdot 1}{a_1 + a_2} R_2 = \frac{a_2}{a_1 + a_2} y$$

b) ~~The~~ 3 function lead to the same result that $y = \min(-a_1, a_2)$. But the network c give the most intuitive explanation. because in the hidden layer, it ~~provides~~ describe 3 different kind of relationship between a_1 and a_2.

~~$a_1 > a_2$~~ $a_1 = a_2$, $a_1 < a_2$, $a_1 > a_2$ respectively.