1

$$\underline{y} = \underline{X}\,\underline{w}$$

$$J(\underline{w}) = (\underline{X}\,\underline{w} - \underline{y})^T(\underline{X}\,\underline{w} - \underline{y}) + \lambda\,\underline{w}^T\underline{w}$$

$$= \left(\underline{w}^T\underline{X}^T\underline{X}\,\underline{w} - 2\underline{w}^T\underline{X}^T\underline{y} - \underline{y}^T\underline{y}\right) + \lambda\,\underline{w}^T\underline{w}$$

use $\dfrac{\partial}{\partial \underline{x}}(\underline{x}^T\underline{A}) = \underline{A}$

$$\frac{\partial}{\partial \underline{x}}(\underline{x}^T\underline{A}\,\underline{x}) = 2\underline{A}\,\underline{x}$$

$$\frac{\partial}{\partial \underline{x}}(\underline{x}^T\underline{x}) = 2\underline{x}$$

$$\frac{\partial J(\underline{w})}{\partial \underline{w}} = 2\underline{X}^T\underline{X}w - 2\underline{y}^T\underline{X} + \lambda \underline{w}$$

$$0 = 2\underline{X}^T\underline{X}\,\underline{w}^* - 2\underline{y}^T\underline{X} + 2\lambda\underline{w}^*$$

$$\underline{w}^* = (\underline{X}^T\underline{X} + \lambda\underline{I})^{-1}\,\underline{y}^T\underline{X}$$

2 $$z_i^{(l)} = \sum_{j=1}^{n_{inputs}} a_i^{(l-1)} w_{ji}^{(l)}$$

$$\mathrm{Var}[z_i^{(l)}] = n_{inputs}\,\mathrm{Var}[a_i^{(l-1)}] \times \mathrm{Var}[w_{ji}^{(l)}]$$

For $\mathrm{Var}[z_i^{(l)}] = \mathrm{Var}[a_i^{(l-1)}]$

require $\mathrm{Var}[w_{ji}^{(l)}] = \frac{1}{n_{inputs}}$.

$$f(x) = \begin{cases} \dfrac{1}{2\alpha} & -\alpha \le x \le \alpha \\ 0 & \text{otherwise} \end{cases}$$

$$E(X) = \int_{-\alpha}^{\alpha} x \, f(x) \, dx$$

$$= \frac{1}{2\alpha} \int_{-\alpha}^{\alpha} x \, dx$$

$$= \frac{1}{4\alpha} \left[x^2 \right]_{-\alpha}^{\alpha} = 0$$

(as expected)

$$E(x^2) = \int_{-\alpha}^{\alpha} x^2 f(x) \, dx$$

$$= \frac{1}{2\alpha} \int_{-\alpha}^{\alpha} x^2 \, dx$$

$$= \frac{1}{6\alpha}\left[x^3\right]_{-\alpha}^{\alpha} = \frac{\alpha^2}{3}$$

$$Var(x) = E(x^2) - E^2(x)$$

$$= \frac{\alpha^2}{3}$$

$$\therefore \quad \alpha = \sqrt{\frac{3}{n_{inputs}}}$$