(*)
GD iter. $x^{n+1} = x^n - \eta \dfrac{\nabla f(x^n)}{\|\nabla f(x^n)\|}$.

Homework

NR iter $x^{n+1} = x^n - H_f^{-1}(x^n)\nabla f(x^n)$.

1. Computation

$H_Q(x^n) \Rightarrow Q(x^n) = \frac{1}{2}$

$$H(p) = -\int p(x)\ln p(x)\,dx$$

invertible

Exercise 4.17.2 Consider the quadratic function $Q(x) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - b\mathbf{x}$, with A nonsingular square matrix of order n.

(a) Find the gradient ∇Q; $\nabla Q = Ax - b$

(b) Write the gradient descent iteration; $x^{n+1} = x^n - \eta \cdot \dfrac{\nabla Q(x^n)}{\|\nabla Q(x^n)\|} = x^n - \eta \cdot \dfrac{Ax^n - b}{\|Ax^n - b\|}$

(c) Find the Hessian H_Q; $H_Q = A$

(d) Write the iteration given by Newton's formula and compute its limit.

$x^{n+1} = x^n - H_Q^{-1}(x^n) \cdot \nabla Q(x^n)$ ✓

Exercise 6.6.10 Let $p(x)$ be the uniform distribution over the interval $[a, b]$. Show that $H(p) = \ln(b - a)$.

$p(x) = \frac{1}{b-a},\ a \le x \le b$ ✓

$H(p) = -\int_a^b \frac{1}{b-a} \cdot \ln \frac{1}{b-a}\,dx = -\frac{1}{b-a} \cdot \ln \frac{1}{b-a} \cdot (b-a) = -\ln \frac{1}{b-a} = \ln(b-a)$ ✓

Exercise 6.6.11 Let $p(x)$ be the one-dimensional normal distribution with mean μ and standard deviation σ. Show that its entropy is $H(p) = \ln(\sigma\sqrt{2\pi e})$.

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{1}{2\sigma^2} \cdot (x-\mu)^2\right]$$

$$H(p) = -\int p(x) \cdot \ln p(x)\, dx$$

$$= -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{1}{2\sigma^2} \cdot (x-\mu)^2\right] \cdot \left\{ -\frac{1}{2} \ln 2\pi\sigma^2 - \frac{1}{2\sigma^2} \cdot (x-\mu)^2 \right\} dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \left(\frac{1}{2} \cdot \ln 2\pi\sigma^2\right) \cdot \exp\left[-\frac{1}{2\sigma^2} \cdot (x-\mu)^2\right] dx$$

$$+ \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{1}{2\sigma^2} (x-\mu)^2\right] \cdot \frac{1}{2\sigma^2} \cdot (x-\mu)^2 \, dx$$

$$= \frac{1}{2} \cdot \ln 2\pi\sigma^2 \cdot \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{1}{2\sigma^2} \cdot (x-\mu)^2\right] dx}_{=1}$$

$$+ \frac{1}{2\sigma^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{1}{2\sigma^2} \cdot (x-\mu)^2\right] \cdot (x^2 - 2\mu x + \mu^2)\, dx$$

Let $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{1}{2\sigma^2} (x-\mu)^2\right]$

$$= \frac{1}{2} \cdot \ln 2\pi\sigma^2 + \frac{1}{2\sigma^2} \cdot \int_{-\infty}^{\infty} x^2 f(x) - 2\mu \cdot x f(x) + \mu^2 f(x)\, dx$$

$$= \frac{1}{2} \ln 2\pi\sigma^2 + \frac{1}{2\sigma^2} \left\{ E(x^2) - 2\mu \cdot E(x) + \mu^2 \right\}$$

$$= \frac{1}{2} \ln 2\pi\sigma^2 + \frac{1}{2\sigma^2} \left\{ \underline{E(x^2) - 2\mu^2 + \mu^2} \right\}$$

$$= \frac{1}{2} \ln 2\pi\sigma^2 + \frac{1}{2} \cdot \frac{1}{\sigma^2} \times V(x)$$

$$= \frac{1}{2} (\ln 2\pi\sigma^2 + \ln e)$$

$$= \frac{1}{2} \cdot \ln 2\pi\sigma^2 \cdot e$$

$$= \ln \sigma \cdot \sqrt{2\pi e}$$

Homework

3. Interpretation

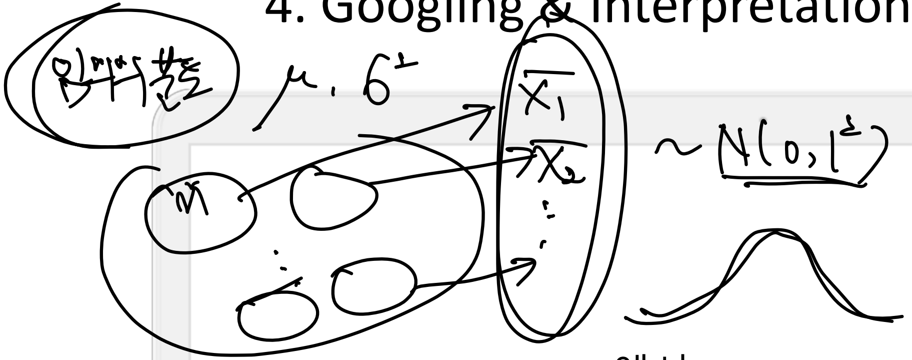
$I \longrightarrow N \longrightarrow O$

$x \qquad a \cdot x + b \qquad ax+b = y$

Exercise 5.10.4 Explain the equivalence between the linear regression algorithm and the learning of a linear neuron.

Homework

4. Googling & Interpretation & Application



Batch Training Section에서, **Central Limit Theorem**에 따라서,
각각의 training을 하는 것보다, Batch training이
분산이 더 작게 나온다는 점에 대해 고민해보아요!