Linear Algebra

1. $x^{(1)} = [a_1, a_2, a_3]$, $x^{(2)} = [a_1, -a_2, a_3]$

If 'θ' be angle betwn $x^{(1)}$ & $x^{(2)}$,

$$\cos\theta = \frac{x^{(1)} \cdot x^{(2)}}{|x^{(1)}| \cdot |x^{(2)}|} = \frac{a_1 \cdot a_1 - a_2 \cdot a_2 + a_3 \cdot a_3}{a_1^2 + a_2^2 + a_3^2} = \frac{a_1^2 - a_2^2 + a_3^2}{a_1^2 + a_2^2 + a_3^2}$$

when $\theta = \pi/2$, $\cos\theta = 0$.

$$\Rightarrow a_1^2 - a_2^2 + a_3^2 = 0.$$

2. Eqn of hyperplane : $\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \dots \theta_d x_d = 0$

Normal to hyperplane : $[\theta_1, \theta_2, \theta_3, \theta_4, \dots, \theta_d]^T = n$

Unit Normal vector : $\frac{[\theta_1, \theta_2, \theta_3, \theta_4 \dots \theta_d]^T}{|n|} = \hat{n}$

$$|n| = \sqrt{\theta_1^2 + \theta_2^2 + \dots \theta_d^2}$$

If d be a vector from a point $(x_1^0, x_2^0, x_3^0, \dots x_d^0) = x^0$ to x then, $\vec{d} = (x_1 - x_1^0, x_2 - x_2^0, \dots, x_d - x_d^0)$

$$\perp^{ar} \text{ distance} = |\hat{n} \cdot \vec{d}|$$

$$= \frac{\theta_1(x_1 - x_1^0) + \theta_2(x_2 - x_2^0) + \dots \theta_d(x_d - x_d^0)}{|n|}$$

$$= \frac{\theta_1 x_1 + \theta_2 x_2 + \dots + \theta_d x_d - (\theta_1 x_1^0 + \theta_2 x_2^0 \dots + x_d^0 \theta_d)}{|n|}$$

$$\perp^{ar} \text{ distance} = \frac{\theta_1 x_1 + \theta_2 x_2 + \dots + \theta_d x_d + \theta_0}{|n|}$$

Probability.

(a) The value of $P_X(x)$ lies in interval $[0,1] \rightarrow$ $\boxed{\text{false}}$

(b) When $a<b$, $\int_a^b P_X(x)\,dx \in [0,1]$ & represents prob. that the value of random variable falls within $[a,b]$ $\rightarrow$ $\boxed{\text{TRUE}}$

(c) $P_X(x)$ is always non-negative. $\boxed{\text{TRUE}}$

(d) The integral $P_X(x)$ from $-\infty$ to ∞ is finite but specific value may vary. $\boxed{\text{FALSE.}}$

(e). $\int_{-\infty}^{\infty} P_X(x)\,dx = 1$ always. $\boxed{\text{TRUE.}}$

5. Univariate Gaussians

$$N(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$\frac{dN}{dx} = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \cdot -\frac{1}{2\sigma^2} \cdot 2(x-\mu)$$

For max point x; $\frac{dN}{dx} = 0 \Rightarrow (x-\mu) = 0$

$\Rightarrow \boxed{x = \mu}$

For Max value of N; $N = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\frac{1}{2}\left(\frac{\mu-\mu}{\sigma}\right)^2}$

$= \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot 1$

$$\boxed{N = \frac{1}{\sqrt{2\pi}\cdot\sigma}}$$

5.c. ~~Univa.~~ $D = \{x_1, x_2, x_3, \dots, x_n\}$.

Since the points are drawn independently,

$$P_X(x_1, x_2, \dots x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2}$$

6 a. Optimization gradient.

$$L(x,\theta) = \log(1 + \exp(-\theta_1 x_1 - \theta_2 x_2))$$

$$\frac{\partial(x,\theta)}{\partial\theta_1} = -\frac{\exp(\theta_1 x_1 - \theta_2 x_2)\, x_1}{1+\exp(-\theta_1 x_1 - \theta_2 x_2)}$$

$$\frac{\partial}{\partial\theta_2}(x,\theta) = \frac{-\exp(\theta_1 x_1 - \theta_2 x_2) \cdot x_2}{1+\exp(-\theta_1 x_1 - \theta_2 x_2)}$$

$$\nabla L(x,\theta) = \frac{-\exp(\theta_1 x_1 - \theta_2 x_2)}{1+\exp(-\theta_1 x_1 - \theta_2 x_2)} \quad (x_1 \hat{i} + x_2 \hat{j}).$$

b. ~~f~~ ϕ be angle from tve x dir,

$$\tan\phi = \frac{x_2}{x_1} \Rightarrow \phi = \tan^{-1}\left(\frac{x_2}{x_1}\right) = \text{dir.}$$

with app. choice of $\pi + \phi$, $\pi - \phi$, $2\pi - \phi$

The value of $L(x,\theta)$ is larger if we evaluate it at

$$\theta' = \theta + \varepsilon \cdot \nabla_\theta L(x,\theta) \quad \text{where } \varepsilon > 0.$$

5.c. ~~Univa~~ $D = \{x_1, x_2, x_3, \ldots, x_n\}$.

Since the points are drawn independently,

$$P_X(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}$$

6 a. Optimization gradient.

$$L(x,\theta) = \log\left(1+\exp\left(-\theta_1 x_1 - \theta_2 x_2\right)\right)$$

$$\frac{\partial (x,\theta)}{\partial \theta_1} = -\frac{\exp(\theta_1 x_1 - \theta_2 x_2)\, x_1}{1+\exp(-\theta_1 x_1 - \theta_2 x_2)}$$

$$\frac{\partial (x,\theta)}{\partial \theta_2} = \frac{-\exp(\theta_1 x_1 - \theta_2 x_2)\cdot x_2}{1+\exp(-\theta_1 x_1 - \theta_2 x_2)}$$

$$\nabla L(x,\theta) = \frac{-\exp(\theta_1 x_1 - \theta_2 x_2)}{1+\exp(-\theta_1 x_1 - \theta_2 x_2)} \left(x_1 \hat{i} + x_2 \hat{j}\right).$$

b. If ϕ be angle from the x dir,

$$\tan\phi = \frac{x_2}{x_1} \Rightarrow \phi = \tan^{-1}\left(\frac{x_2}{x_1}\right) = \text{dir.}$$

with app. choice of $\pi+\phi$, $\pi-\phi$, $2\pi-\phi$

The value of $L(x,\theta)$ is larger if we evaluate it at

$$\theta' = \theta + \varepsilon \cdot \nabla_\theta L(x,\theta) \quad \text{where } \varepsilon > 0.$$

f.

$$A^T(AB - C) = 0$$

$$\Rightarrow \quad A^T AB - A^T C = 0$$

$$\Rightarrow \quad A^T AB = A^T C$$

$$\Rightarrow \quad (A^T)^{-1} A^T AB = A^{T^{-1}} A^T C$$

$$\Rightarrow \quad AB = C$$

$$\Rightarrow \quad A^{-1}AB = A^{-1}C$$

$$\boxed{B = A^{-1}C}$$