

$$(a) \mathcal{L}(s(x))_{x \in G}(\lambda_1, \lambda_2, \lambda_3) = - \sum_{x \in G} \exp(s(x)) s(x) \delta + \lambda_1 \left(\sum_{x \in G} \exp(s(x)) \delta - 1 \right) \\ + \lambda_2 \left(\sum_{x \in G} \exp(s(x)) x \delta \right) + \lambda_3 \left(\sum_{x \in G} \exp(s(x)) x^2 \delta - \delta^{-1} \right)$$

(b) Let $\hat{x} \in G$.

$$0 \stackrel{!}{=} \frac{\partial}{\partial s(\hat{x})} \mathcal{L}(s(x))_{x \in G}(\lambda_1, \lambda_2, \lambda_3) \\ = - \exp(s(\hat{x})) (1 + s(\hat{x})) \delta + \lambda_1 \exp(s(\hat{x})) \delta + \lambda_2 \exp(s(\hat{x})) \hat{x} \delta + \lambda_3 \exp(s(\hat{x})) \hat{x}^2 \delta \\ \Leftrightarrow 0 = -(1 + s(\hat{x})) + \lambda_1 + \lambda_2 \hat{x} + \lambda_3 \hat{x}^2 \quad (I) \\ \Leftrightarrow s(\hat{x}) = \lambda_3 \hat{x}^2 + \lambda_2 \hat{x} + \lambda_1 - 1 \quad (II)$$

The number λ_3 has to be negative. If this was not the case, then, by equation (I), s would either be constant or tend to $+\infty$ for $\hat{x} \rightarrow +\infty$ or for $\hat{x} \rightarrow -\infty$. Since G is a regular partition of \mathbb{R} , there have to exist sequences in G tending to $\pm\infty$. Now if s is constant or unbounded, this contradicts $0 \stackrel{!}{=} \frac{\partial \mathcal{L}}{\partial \lambda_1} = \sum_{x \in G} \exp(s(x)) \delta - 1$ (II), the series would diverge to $+\infty$. Hence, $\lambda_3 < 0$.

Rewriting (I) gives

$$s(\hat{x}) = \lambda_3 \left(\hat{x}^2 + \frac{\lambda_2}{\lambda_3} \hat{x} \right) + \lambda_1 - 1 \\ = \lambda_3 \left(\hat{x} + \frac{1}{2} \cdot \frac{\lambda_2}{\lambda_3} \right)^2 - \frac{1}{4} \cdot \frac{\lambda_2^2}{\lambda_3} + \lambda_1 - 1 \\ = -\frac{1}{2\theta^2} (\hat{x} - \mu)^2 + r,$$

setting $\theta^2 := \sqrt{-\frac{1}{2\lambda_3}}$, $\mu := -\frac{1}{2} \cdot \frac{\lambda_2}{\lambda_3}$, $r := -\frac{1}{4} \cdot \frac{\lambda_2^2}{\lambda_3} + \lambda_1 - 1$.

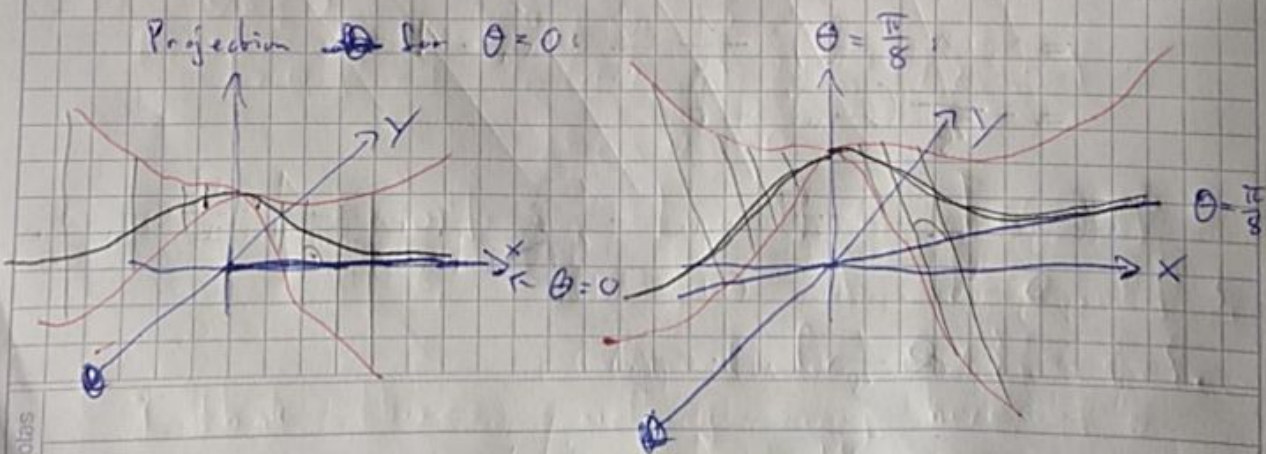
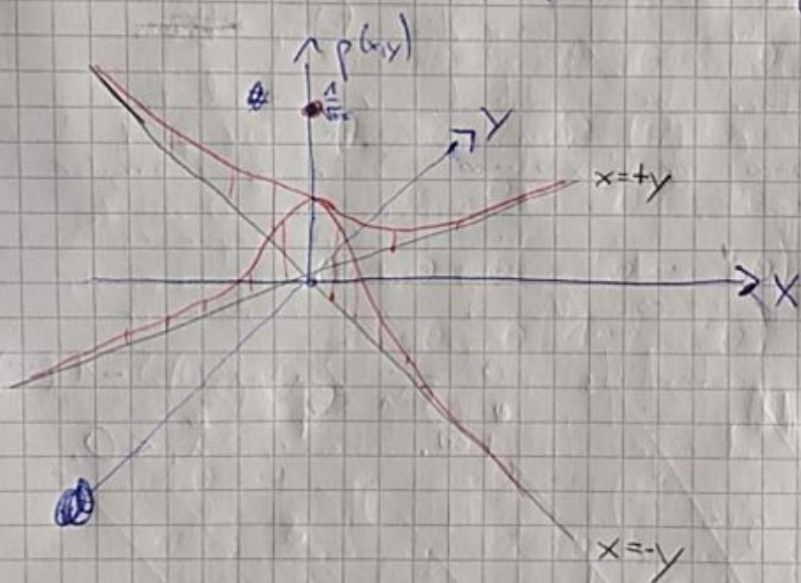
Plugging into equation (II) yields

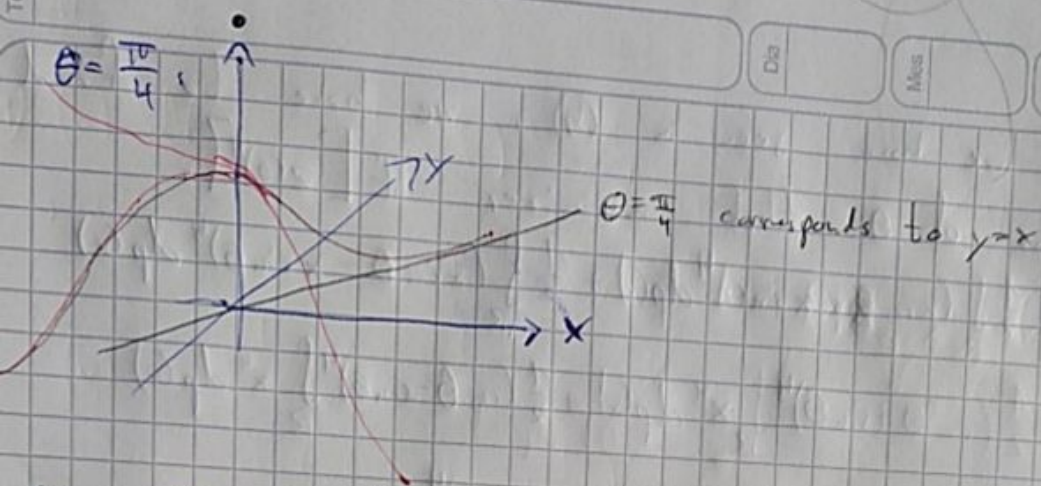
$$1 = \sum_{x \in G} \exp\left(-\frac{1}{2\theta^2} (\hat{x} - \mu)^2\right) \cdot \exp(r) \delta,$$

hence $r = -\ln \left(\sum_{x \in G} \exp\left(-\frac{1}{2\theta^2} (\hat{x} - \mu)^2\right) \delta \right).$

At this point we can already see that $p(x,y)$ has a normal distribution with mean μ and variance σ^2 . By construction, taking the derivatives of ℓ wrt. μ, σ^2 will yield $\mu=0, \sigma^2=6^2$.

- (2) (a) The joint pdf consists of the pdf of a standard normal distribution scaled by the factor $\frac{1}{2}$ along the ~~axis~~ axes $y=x$ and $y=-x$ with a jump to $\frac{1}{\sqrt{2\pi}}$ in 0.
- $$p(x,y) = p(x)p(y|x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \cdot \begin{cases} 1, & \text{if } x=y=0 \\ \frac{1}{2}, & \text{if } x=y \neq 0 \text{ or } x=-y \\ 0, & \text{else} \end{cases}$$





$$(b) \cdot E[z(\theta)] = \int_{\mathbb{R}^2} (x \cos \theta + y \sin \theta) d\phi(x, y)$$

$$\int_{\mathbb{R}^2} (x \cos \theta + y \sin \theta) d\phi(x, y)$$

$$= \int_{\mathbb{R}} (x \cos \theta + x \sin \theta) \cdot \frac{1}{2} d\phi(x) + \int_{\mathbb{R}} (x \cos \theta - x \sin \theta) \cdot \frac{1}{2} d\phi(x)$$

$$= \int_{\mathbb{R}} x \cos \theta d\phi(x)$$

$$= \cos \theta \cdot E[x]$$

$$= 0$$

$$\cdot V[z(\theta)] = \int_{\mathbb{R}^2} (x \cos \theta + y \sin \theta)^2 d\phi(x, y)$$

$$= \int_{\mathbb{R}} \frac{1}{2} (x \cos \theta + x \sin \theta)^2 d\phi(x) + \int_{\mathbb{R}} \frac{1}{2} (x \cos \theta - x \sin \theta)^2 d\phi(x)$$

$$= \int_{\mathbb{R}} \frac{1}{2} x^2 ((\cos \theta + \sin \theta)^2 + (\cos \theta - \sin \theta)^2) d\phi(x)$$

$$= \int_{\mathbb{R}} x^2 (\cos^2 \theta + \sin^2 \theta) d\phi(x)$$

$$= (\cos^2 \theta + \sin^2 \theta) \cdot V[x]$$

$$= \cos^2 \theta + \sin^2 \theta$$

$$= 1$$

Hence, the variance is constant and ~~is~~ maximized by all $\theta \in [0, 2\pi]$

By what we have seen in (b), we have to find the value of θ that maximizes $E[z(\theta)^4]$.

$$\begin{aligned} E[z(\theta)^4] &= \frac{1}{2} \int_{\mathbb{R}} (x \cos \theta + x \sin \theta)^4 d\phi(x) + \frac{1}{2} \int_{\mathbb{R}} (x \cos \theta - x \sin \theta)^4 d\phi(x) \\ &= \frac{1}{2} (\cos \theta + \sin \theta)^4 \int_{\mathbb{R}} x^4 d\phi(x) + \frac{1}{2} (\cos \theta - \sin \theta)^4 \int_{\mathbb{R}} x^4 d\phi(x) \\ &= \frac{3}{2} (\cos \theta + \sin \theta)^4 + \frac{3}{2} (\cos \theta - \sin \theta)^4, \end{aligned}$$

where we used that the 4th moment of a standard normal distribution is 3.

$$\begin{aligned} \frac{d}{d\theta} E[z(\theta)^4] &= 6(\cos \theta + \sin \theta)^3 (-\sin \theta + \cos \theta) + 6(\cos \theta - \sin \theta)^3 (-\sin \theta - \cos \theta) \\ &= 6(\cos^3 \theta - 3\cos^2 \theta \sin \theta + 3\cos \theta \sin^2 \theta + \sin^3 \theta)(-\sin \theta + \cos \theta) \\ &\quad - 6(\cos^3 \theta - 3\cos^2 \theta \sin \theta + 3\cos \theta \sin^2 \theta - \sin^3 \theta)(\sin \theta + \cos \theta) \\ &= -12\cos^3 \theta \sin \theta + 36\cos^3 \theta \sin \theta - 36\cos \theta \sin^3 \theta + 12\cos \theta \sin^3 \theta \\ &= 12\cos \theta \sin \theta (-\cos^2 \theta + 3\cos^2 \theta - 3\sin^2 \theta + \sin^2 \theta) \\ &= 24\cos \theta \sin \theta (\cos^2 \theta - \sin^2 \theta) \\ &= 24\cos \theta \sin \theta (1 - 2\sin^2 \theta) \\ &\stackrel{!}{=} 0 \end{aligned}$$

$$\Leftrightarrow \cos \theta = 0 \quad \text{or} \quad \sin \theta = 0 \quad \text{or} \quad \sin \theta = \frac{1}{\sqrt{2}}$$

$$\Leftrightarrow \theta \in \left\{ \frac{\pi}{2}, \frac{3\pi}{2} \right\} \quad \text{or} \quad \theta \in \{0, \pi\} \quad \text{or} \quad \theta \in \left\{ \frac{\pi}{4}, \frac{3\pi}{4} \right\}$$

$$\Leftrightarrow \theta \in \left\{ 0, \frac{1}{4}\pi, \frac{1}{2}\pi, \frac{3}{4}\pi, \pi, \frac{3}{2}\pi \right\}$$

$$\frac{d^2}{d\theta^2} E[z(\theta)^4] = -24 \sin \theta (1 - 2\sin^2 \theta) + 24 \cos^2 \theta (1 - 2\sin^2 \theta)$$

Since $E[z(0)^4] = 3$, $E[z(\frac{1}{4}\pi)^4] = 6$, $E[z(\frac{1}{2}\pi)^4] = 3$, $E[z(\frac{3}{4}\pi)^4] = 6$, $E[z(\pi)^4] = 3$ and $E[z(\frac{3}{2}\pi)^4] = 3$, the maximal value is attained for $\theta \in \left\{ \frac{1}{4}\pi, \frac{3}{4}\pi \right\}$.

I think there's something wrong here, if θ is a solution, then also $\theta + \pi$ should be one.

3 (a) • $E[\bar{w}^T x] = \sum_{k=1}^d w_k \underbrace{E[x_k]}_{=0} = 0$, since $E[x] = 0 \in \mathbb{R}^d$.

• $V[\bar{w}^T x] = E[(\bar{w}^T x)^2] = E[\bar{w}^T x x^T \bar{w}] = \bar{w}^T E[x x^T] \bar{w} = \bar{w}^T \bar{w} = \|\bar{w}\|_2^2 =$

(b) Using (a), the Lagrangian can be written as

$$\mathcal{L}(\bar{w}, \lambda) = E[(\bar{w}^T x)^2] + \lambda (\|\bar{w}\|_2^2 - 1).$$

Supposing we can interchange integration and differentiation, we get

$$0 = \frac{\partial}{\partial \bar{w}} \mathcal{L}(\bar{w}, \lambda) = E[4(\bar{w}^T x)^2 x] + 2\lambda \bar{w}$$

$$\Rightarrow \frac{\partial}{\partial \bar{w}} \mathcal{L}(\bar{w}, \lambda) = E[(\bar{w}^T x)^2 x],$$

which is a symmetric problem for the use in the exercise since λ can be removed by multiplying by \bar{w}^T from the left and using $\|\bar{w}\|_2^2 = 1$.

λ can be found by multiplication of \bar{w}^T from the left:

$$\lambda \bar{w}^T \bar{w} = E[(\bar{w}^T x)^2 \bar{w}^T x] = E[(\bar{w}^T x)^2]$$

$$(c) F(\bar{w}) := E[(\bar{w}^T x)^2 x] - \lambda \bar{w}, \text{ i.e., } F(\bar{w}) = E[(\bar{w}^T x)^2 x] - \lambda \bar{w};$$

$$\Rightarrow \frac{\partial F(\bar{w})}{\partial \bar{w}_j} = E[2(\bar{w}^T x)^2 x_j x_j] - \lambda \delta_{ij}, \text{ hence we can write}$$

$$J(\bar{w}) = E[2(\bar{w}^T x)^2 x x^T] - \lambda \text{Id} = E[2(\bar{w}^T x)^2 x x^T - \lambda \text{Id}]$$

Plugging into the formula $\bar{w}^T = \bar{w} - J(\bar{w})^{-1} F(\bar{w})$ yields the formula.

$$(d) \bar{w}^T = \bar{w} - (E[2 x x^T (\bar{w}^T x)^2] - \lambda \text{Id})^{-1} (E[x (\bar{w}^T x)^2] - \lambda \bar{w})$$

$$= \bar{w} - (2 \underbrace{E[x x^T]}_{=\text{Id}} \underbrace{E[(\bar{w}^T x)^2]}_{=\text{Id by (a)}} - \lambda \text{Id})^{-1} (E[x (\bar{w}^T x)^2] - \lambda \bar{w})$$

$$= \bar{w} - \frac{1}{2-\lambda} \underbrace{\text{Id}^{-1}}_{=\text{Id}} (E[x (\bar{w}^T x)^2] - \lambda \bar{w})$$

$$= -\frac{1}{2-\lambda} E[x (\bar{w}^T x)^2] + \frac{(1+\frac{\lambda}{2-\lambda})}{2-\lambda} \bar{w} = -\frac{1}{2-\lambda} (E[x (\bar{w}^T x)^2] - 3\bar{w})$$