Total $\boxed{80/100}$ (B+)

Problem 1

[Your potential is A, please work it!]

(a) $M_X(\lambda) = E[e^{\lambda X}]$, $e^{\lambda X} = \sum_{k=0}^{\infty} \frac{(\lambda X)^k}{k!}$ (power series)

$$M_X(\lambda) = E\left[\sum_{k=0}^{\infty} \frac{\lambda^k X^k}{k!}\right] = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} E[X^k]$$

By Markov's inequality

$$P(X > u) = P(e^{\lambda X} > e^{\lambda u}) \le e^{-\lambda u} M_X(\lambda)$$

taking the infimum over $\lambda > 0$

$$P(X > u) \le \inf_{\lambda>0} e^{-\lambda u} M_X(\lambda)$$

expanding the MGF in the bound:

$$\inf_{\lambda>0} e^{-\lambda u} M_X(\lambda) = \inf_{\lambda>0} e^{-\lambda u} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} E[X^k]$$

— what you're choosing here is k, not λ. (Why?)

choosing $\lambda = \frac{k}{u}$, we simplify $e^{-\lambda u} = e^{-k}$

and since $k!$ is in the denominator, for large k

$$\Rightarrow e^{-k} \frac{k^k}{k!} u^{-k} E[X^k] \approx E[X^k] u^{-k}$$

$k! \approx \sqrt{2\pi k}\left(\frac{k}{e}\right)^k$

$\Rightarrow \frac{k^k}{k!} \approx e^k k^{-k}$

$$\Rightarrow \inf_{\lambda>0} M_X(\lambda) e^{-\lambda u} \geq \inf_{k \in \mathbb{Z}_+} E[X^k] u^{-k}$$

$\boxed{\text{Great!}}$

$+$.

This is not what I asked to prove - the result is weaker - but your argument is basically correct → bonus points for nonstandard thinking!

(b)

MGF: $M_X(\lambda) = E[e^{\lambda X}]$

since X is symmetric, $E[e^{\lambda X}] = E[e^{-\lambda X}]$

$\Rightarrow M_X(\lambda) = E\left[\frac{e^{\lambda X} + e^{-\lambda X}}{2}\right]$

power series: $e^{\lambda X} + e^{-\lambda X} = 2\sum_{k=0}^{\infty} \frac{\lambda^{2k} X^{2k}}{(2k)!}$

taking expectations; $M_X(\lambda) = \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{(2k)!} E[X^{2k}]$

since X is symmetric, all odd moments vanish ($E[X^k]=0$ for all odd k) Yup

From Markov's inequality and Chernoff's method

$$P(X>u) = P(e^{\lambda X} > e^{\lambda u}) \leq e^{-\lambda u} M_X(\lambda)$$

taking infimum over $\lambda > 0$

$$P(X>u) \leq \inf e^{-\lambda u} M_X(\lambda)$$

Since X is symmetric, using 2 sided bound:

$$P(|X| \geq u) = P(X>u) + P(X<-u) = 2P(X>u)$$

$$P(|X|>u) \leq 2\inf_{\lambda>0} e^{-\lambda u} M_X(\lambda),$$

moment-based bound: $P(|X|>u) \leq \inf \frac{E[X^{2k}]}{u^{2k}}$ (sharper alternative)

expanding the MGF bound

$$e^{-\lambda u} M_X(\lambda) = e^{-\lambda u} \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{(2k)!} E[X^{2k}]$$

choosing $\lambda = \frac{2k}{u}$, $\frac{(2k)^{2k}}{(2k)!} \approx e^{2k}(2k)^{-2k}$

symmetry

taking infimum over k

$$\Rightarrow \inf_{\lambda>0} M_X(\lambda) e^{-\lambda u} \geq \frac{1}{2} \inf_{k \in \mathbb{Z}_+} E[X^{2k}] u^{-2k}$$

Problem 2

(a) X : random variable

applying Young's inequality :

Not so easy:

$$|a^T b| \leq \|a\|_p \|b\|_q$$

$$\left.\begin{array}{l}\|t\|_p = t \\ \|X\|_q = |X|\end{array}\right\}$$ for any $t \in \mathbb{R}$, $X \in \mathbb{R}$

$$\cancel{tX = (t \cdot X) \leq \|t\|_p \|X\|_q}$$

$$\cancel{E[e^{tX}] \leq E[e^{\|t\|_p \|X\|_q}]}$$

$$\cancel{K_X(t) = \log E[e^{tX}] \leq \log E[e^{\|t\|_p \|X\|_q}]}$$

$$\cancel{\Rightarrow K_X(t) \text{ convex}}$$

other way of solving using Jensen's inequality

$$K_X(\lambda t_1 + (1-\lambda)t_2) = \log E[e^{(\lambda t_1 + (1-\lambda)t_2)X}]$$

$$e^{(\lambda t_1 + (1-\lambda)t_2)X} \leq \lambda e^{t_1 X} + (1-\lambda) e^{t_2 X}$$

wrong implication

$$\log E[e^{\lambda t_1 X + (1-\lambda) t_2 X}] \leq \lambda \log E[e^{t_1 X}] + (1-\lambda) \log E[e^{t_2 X}]$$

(log(·) is concave, not convex!)

$$K_X(\lambda t_1 + (1-\lambda)t_2) \leq \lambda K_X(t_1) + (1-\lambda) K_X(t_2)$$

$\ominus$

$$\Rightarrow K_X(t) \text{ convex}$$

Problem 3

(a) $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ is the pdf of $N(0,1)$

$$\Phi(u) = \int_u^\infty \phi(t)\,dt$$

substituting $\phi(t) = \frac{-\phi'(t)}{t}$ and using integration by p

$u = \frac{1}{t} \Rightarrow du = -\frac{1}{t^2}\,dt$

$dv = \phi'(t) \Rightarrow v = -\phi(t)$

$$\Rightarrow \Phi(u) = \int_u^\infty \phi(t)\,dt = \int_u^\infty -\frac{\phi'(t)}{t}\,dt = \left[\frac{-\phi(t)}{t}\right]_u^\infty - \int_u^\infty \frac{\phi(t)}{t^2}\,dt$$

as $t \to \infty$, $\frac{-\phi(t)}{t} = 0$, we get

$$\Phi(u) = \frac{\phi(u)}{u} - \underbrace{\int_u^\infty \frac{\phi(t)}{t^2}\,dt}_{>0}$$

$$\Rightarrow \Phi(u) \leq \frac{\phi(u)}{u}$$

applying integration by parts again

$$\frac{\Phi(u)}{u} - \int_u^\infty \frac{-\phi'(t)}{t^3} = \frac{\phi(u)}{u} + \left[\frac{\phi(t)}{t^3}\right]_u^\infty - \int_u^\infty \frac{-3\phi(t)}{t^4}\,dt$$

$$= \frac{\phi(u)}{u} - \frac{\phi(u)}{u^3} + \int_u^\infty \frac{3\phi(t)}{t^4}\,dt$$

since $\int_u^\infty \frac{3\phi(t)}{t^4}\,dt > 0$

$$\Rightarrow \phi(u)\left(\frac{1}{u} - \frac{1}{u^3}\right) < P[u \geq u] - \Phi(u)$$

$$\Phi(u) \geq \left(\frac{1}{u} - \frac{1}{u^3}\right)\phi(u)$$

+

(b) Applying the same trick again as in part (a) to $\int_u^\infty \frac{3\phi(t)}{t^4}\,dt$

$$\Rightarrow \int_u^\infty \frac{-3\phi'(t)}{t^5}\,dt = \left[-\frac{3\phi(t)}{t^5}\right]_u^\infty - \int_u^\infty \frac{15\phi(t)}{t^6}\,dt$$

$$= \frac{3\phi(u)}{u^5} - \int_u^\infty \frac{15\phi(t)}{t^6}\,dt$$

since $-\int_z^\infty \frac{15\phi(t)}{t^6}\,dt < 0$

$$\Rightarrow \Phi(u) \leq \Phi(u)\left(\frac{1}{u} - \frac{1}{u^3} + \frac{3}{u^5}\right)$$

+

3.2) $\Phi(u) = \int_{-\infty}^u \phi(t)\,dt$

where $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$

since the normal distribution is symmetric around 0,

$$\Phi(u) = \frac{1}{2} + \int_0^u \phi(t)\,dt$$

$$\Rightarrow \frac{1}{2} - \Phi(u) = \int_u^\infty \phi(t)\,dt$$

Gaussian tail probability

$$\frac{1}{2} - \Phi(u) = \int_u^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\,dt$$

substitute: $t = u + x$ $\quad dt = dx$

$$\frac{1}{2} - \Phi(u) = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-(u+x)^2/2}\,dx$$

$$\frac{1}{2} - \Phi(u) = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-ux} e^{-x^2/2}\,dx$$

when $t = u$, $x = 0$
for integral bounds

$$\frac{1}{2} - \Phi(u) = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \sum_{k=0}^{\infty} \frac{(-u)^k}{k!} \underbrace{\int_0^\infty x^k e^{-x^2/2} dx}$$

gamma $\quad 2^{(k-1)/2} < \frac{(k+1)}{2}$

for odd k: $\Gamma(n+1) = n!$ $\downarrow$

for $k = 2m+1 \implies \Gamma\left(\frac{2m+2}{2}\right) = \Gamma(m+1) = m!$

$$\implies \int_0^\infty x^k e^{-x^2/2} dx = 2^{\frac{(k-1)}{2}} \frac{k!}{2^{(k+1)/2}(k+1)} = \frac{k!}{2^{k/2}(k+1)}$$

substitute back

$$\implies \frac{1}{2} - \Phi(u) = \frac{1}{\sqrt{2\pi}} \sum_{k=0}^{\infty} \frac{(-1)^k u^{2k+1}}{2^k k! (2k+1)}$$

You need equality, not inequality

$\ominus$

Problem 4

(a) $I = \mathbb{1}_{\{X \geq (1-t)E[X]\}}$ → indicator function

$XI \geq (1-t)E[X]I$

$E[XI] \geq (1-t)E[X]\,E[I]$

these inequalities are correct but are not the ones you need here, as they lead to the upper bound on P.

since I indicator function

$E[XI] = E[X \mid X \geq (1-t)E[X]]\,P(X \geq (1-t)E[X])$

$E[XI] \geq (1-t)E[X]\,P(X \geq (1-t)E[X])$?

By Cauchy-Schwarz: $\boxed{E[XI]^2 \leq E[X^2]\,\underbrace{E[I]}}$

this is correct $E[I^2] = E[I]$ for indicator function

since $E[I] = P(X \geq (1-t)E[X])$

$E[XI]^2 \leq E[X^2]\,P(X \geq (1-t)E[X])$

$\Rightarrow \boxed{((1-t)E[X]\,P(X \geq (1-t)E[X]))^2 \leq E[X^2]\,P(X \geq (1-t)E[X])}$

dividing both sides by $E[X^2]$ ⇓ that's not what you get

$(1-t)^2 \dfrac{(E[X])^{2-1}}{E[X^2]} \cancel{\leq} \geq P(X \geq (1-t)E[X])$

let set $t = 1-t$

$P(X \geq (1-t)E[X]) \geq t^2 \dfrac{(E[X])^2}{E[X^2]}$

-4

(See my solution – or come to discuss in the office hrs)

(b) $\text{Var}(X) = E[X^2] - (E[X])^2$

$\Rightarrow E[X^2] = \text{Var}(X) + (E[X])^2$

we know, $P(X \geq (1-t)E[X]) \geq t^2 \dfrac{(E[X])^2}{E[X^2]}$

$$\Rightarrow P(X \geq (1-t)E[X]) \geq t^2 \frac{(E[X])^2}{(E[X])^2 + \text{Var}(X)}$$

$$(E[X])^2 + \text{Var}[X] = t^2 (E[X])^2 + \text{Var}(X) + (1-t^2)(E[X])^2$$

taking the worst case scenario where the variance ~~expression has the largest impact~~, we bound:

$$(E[X])^2 + \text{Var}[X] = t^2(E[X])^2 + \text{Var}(X)$$

$$\Rightarrow P(X \geq (1-t)E[X]) \geq t^2 \frac{(E[X])^2}{t^2(E[X])^2 + \text{Var}[X]}$$

an example is the Bernoulli random variable:

$$X = \begin{cases} a & \text{w/ probability } p \\ 0 & \text{w/ probability } 1-p \end{cases}$$

here, $E[X] = pa$, $E[X^2] = pa^2$, $\text{Var}[X] = p(1-p)a^2$

Yes, in the limit $t \to 0$, (b) reduces to (a).
But for any $t > 0$ (b) is stronger.

$\boxed{-}$

(c) From Paley-Zygmund:

$$P(X \geq (1-t)E[X]) \geq \frac{(E[XI])^2}{E[X^2]\,P(X \geq (1-t)E[X])}$$

Hölder's inequality:

$$E[UV] \leq (E[|U|^p]^{1/p})\,(E[|V|^q])^{1/q}$$

$U = X$, $V = I = \mathbb{1}_{\{X \geq (1-t)E[X]\}}$

$\Rightarrow E[V^q] = P(X \geq (1-t)E[X]$

applying Hölder: $E[XI] \leq (E[|X|^p])^{1/p}\,(P(X \geq (1-t)E[X]))^{1/q}$

$\Rightarrow$ raising both sides to p:

$$E[XI]^p \leq E[|X|^p].(P(X \geq (1-t)E[X]))^{p/q}$$

since $p/q = \frac{p-1}{\cancel{q}}$

$$\Rightarrow E[XI]^p \leq E[|X|^p].(P(X \geq (1-t)E[X]))^{\frac{p-1}{\cancel{p}}}$$

using the bound: $E[XI] \geq (1-t)E[X^p]\,P(X \geq (1-t)E[X])$:

$$((1-t)E[X]\,P(X \geq (1-t)E[X]))^p \leq E[|X|^p].(P(X \geq (1-t)E[X]))^{\frac{p-1}{\cancel{p}}}$$

dividing both sides by E[

(+.)

Modulo the $\frac{p-1}{p}$ instead of $p-1$, this would work if (a) did. (You need to adjust solution in the same way as in (a).) So, I didn't penalize here.

Problem 5

(a) For $X \sim \chi_2^2$ $\Rightarrow$ MGF : $M_2(t) = E[e^{tX}]$

$$= \int_0^\infty e^{tx} f_X(x)\, dx$$

with $X = Z_1^2 + Z_2^2$ and $Z_i \sim N(0,1)$

chi-square density function for 2 degrees of freedom: $f_X(x) = \frac{1}{2} e^{-x/2}$, $x > 0$

$$\Rightarrow M_2(t) = \int_0^\infty e^{tx} \cdot \frac{1}{2} e^{-x/2}\, dx$$

$$= \int_0^\infty e^{x(t-1/2)}\, dx$$

the integral above converges if $t < 1/2$
since $\int_0 e^{-\lambda x} = \frac{1}{\lambda}$ for $\lambda > 0$

$\Rightarrow$ in this case, $\lambda = \frac{1}{2} - t$

$$\Rightarrow M_2(t) = \frac{1}{2} \cdot \frac{1}{\frac{1}{2} - t} = \frac{1}{1-2t} \quad , t < 1/2$$ +

using polar coordinates

since $X = Z_1^2 + Z_2^2$, $Z_1 = r\cos\theta$, $Z_2 = r\sin\theta$

pdf : $f_{Z_1,Z_2}(z_1, z_2) = \frac{1}{2\pi} e^{-(z_1^2+z_2^2)/2}$

the Jacobian of the transformation is r

$$\Rightarrow f_{r,\theta}(r,\theta) = \frac{1}{2\pi} e^{-r^2/2} \cdot r$$

$$X = r^2 \Rightarrow M_2(t) = \int_0^\infty e^{tr^2} \underbrace{\left(\int_0^{2\pi} \frac{1}{2\pi} d\theta\right)}_{\text{integral over } \theta = 1} r e^{-r^2/2} dr$$

$$\Rightarrow M_2(t) = \int_0^\infty e^{tr^2} r e^{-r^2/2} dr$$

gamma function: $\int_0^\infty x^{v-1} e^{-\lambda x^p} dx = \frac{\Gamma(v/p)}{p \lambda^{v/p}}$, $\lambda > 0$

$p = 2$, $v = 2$, $\lambda = \frac{1}{2} - t$

$$\Rightarrow \int_0^\infty r e^{-(1/2 - t) r^2} = \frac{1}{1/2 - t}$$

$$\Rightarrow M_2(t) = \frac{1}{1-2t} \quad , t < 1/2$$

$$M_{2d}(t) = \prod_{i=1}^{2d} M_{z_i^2}(t)$$

$$M_{z_i^2}(t) = (1 - 2t)^{-1/2} \quad , t < 1/2$$

since we have 2d independent square normal variables,

$$M_{2d}(t) = (1-2t)^{-2d/2}$$

$$M_{2d}(t) = \frac{1}{(1-2t)^d} \quad , t < 1/2$$

+

(b)

$X \sim \chi^2_{2d}$, chernoff's: $P(X>x) = P(e^{tX} > e^{tx})$
$\leq E[e^{tX}]e^{-tx}$

$$P(X>x) \leq \inf_{t<\frac{1}{2}} \frac{e^{-tx}}{(1-2t)^d}$$

from (a), $M_{2d}(t) = \frac{1}{(1-2t)^d}$, $t < \frac{1}{2}$

$$\Rightarrow \quad P(X>x) \leq \inf_{t<\frac{1}{2}} \frac{e^{-tx}}{(1-2t)^d}$$

$$\log P(X>x) \leq \inf_{t<\frac{1}{2}} (-tx - d\log(1-2t))$$

to find opt t, $\frac{\partial}{\partial t} = -x + \frac{2d}{1-2t} = 0$

$$\Rightarrow t = \frac{x-2d}{2x}$$

$$\log P(X>x) = d\log\left(\frac{x}{2d}\right) - \frac{x-2d}{2}$$

$$P(X>x) = \exp\left(d\log\left(\frac{x}{2d}\right) - \frac{x-2d}{2}\right)$$

$z = x - 2d$, $x = 2d + z$

$$\Rightarrow P(X-2d > z) = \exp\left(d\log\left(\frac{2d+z}{2d}\right) - \frac{z}{2}\right)$$

+

I asked to prove this first ...

Bonus

(c) hint $\Rightarrow$ $\boxed{\log(1+u) \leq u - \frac{1}{4}\min\{u, u^2\}}$, $\forall u \geq 0$

from (b) $\Rightarrow$ $P(X-2d > z) \leq \exp\left(d\log\left(\frac{2d+z}{2d}\right) - \frac{z}{2}\right)$

let's set: $u = z/2d$ $\Rightarrow$ $1+u = \frac{2d+z}{2d}$

$du = \frac{z}{2}$

$\Rightarrow$ $P(X-2d > z) \leq \exp\left(d\log(1+u) - \frac{z}{2}\right)$

$\Rightarrow$ $P(X-2d > z) \leq \exp(d\log(1+u) - du)$

$d\log(1+u) \leq d\left(u - \frac{1}{4}\min\{u, u^2\}\right)$

expanding and rearranging

$d\log(1+u) - du \leq -\frac{d}{4}\min\{u, u^2\}$

plugging in the expression for u : $u = z/2d$

$\Rightarrow$ $P(X-2d > z) \leq \exp\left(-\frac{d}{4}\min\left\{\frac{z}{2d}, \frac{z^2}{4d^2}\right\}\right)$

$P(X-2d > z) \leq \exp\left(-\min\left\{\frac{z}{8}, \frac{z^2}{16d}\right\}\right)$

case 1: if $z > 2d$ $\Rightarrow$ $P(X-2d > z) \leq \exp\left(-\frac{z}{8}\right)$

case 2: if $0 \leq z \leq 2d$ $\Rightarrow$ $P(X-2d > z) \leq \exp\left(-\frac{z^2}{16d}\right)$

$$\Rightarrow P(X-2d > z) \leq \begin{cases} \exp\left(-\frac{z^2}{16d}\right), & 0 \leq z \leq 2d \\ \exp\left(-\frac{z}{8}\right), & z > 2d \end{cases}$$

using c.ii

setting $P(X - 2d > z) = \delta$

$$\Rightarrow \delta = \exp\left(-\min\left\{\frac{z^2}{16d}, \frac{z}{8}\right\}\right)$$

$$\log \delta = -\min\left\{\frac{z^2}{16d}, \frac{z}{8}\right\}$$

$$\min\left\{\frac{z^2}{16d}, \frac{z}{8}\right\} = -\log \delta$$

for $0 \le z \le 2d \Rightarrow \frac{z^2}{16d} = -\log \delta$

$$z = \sqrt{16 d \log 1/\delta}$$

for $z > 2d \Rightarrow \frac{z}{8} = -\log \delta$

$$z = 8 \log 1/\delta$$

$$z \le \max\left\{\sqrt{16 d \log 1/\delta}, 8 \log 1/\delta\right\}$$

$$z \le \sqrt{16 d \log 1/\delta} + 8 \log 1/\delta$$

$$z \le \sqrt{C d \log 1/\delta} + c \log 1/\delta$$

$C = 16$, $c = 8$

$$\Rightarrow z \le \sqrt{16 d \log 1/\delta} + 8 \log 1/\delta$$

f.

Problem 6

(a) $\hat{\mu} = sX$

since $X \sim \mathcal{N}(\mu, I_d) \Rightarrow X = \mu + Z$

and $Z \sim \mathcal{N}(0, I_d)$

$\Rightarrow \hat{\mu} = sX = s(\mu + Z) = s\mu + sZ$

the risk is given by:

$$E_\mu \|sX - \mu\|^2 = E_\mu \|s\mu + sZ - \mu\|^2$$

$$= E_\mu \|(s-1)\mu + sZ\|^2$$

(using: $E\|A+B\|^2 = \|A\|^2 + E\|B\|^2 + 2E\langle A, B\rangle$

$$= (s-1)^2\|\mu\|^2 + s^2 E\|Z\|^2 + 2(s-1)s E\langle \mu, Z\rangle$$

since Z is mean-zero and independent of μ

$\Rightarrow E\langle \mu, Z\rangle = 0 \Rightarrow$ last term vanishes

since $Z \sim \mathcal{N}(0, I_d)$ and $E\|Z\|^2 = d$

$\Rightarrow \text{Risk}(\hat{\mu}) = (s-1)^2\|\mu\|^2 + s^2 d$

case 1 : $s>1 \Rightarrow$ the risk function is strictly increasing in s for $s>1$. choosing $s=1$ gives a lower risk.

$\Rightarrow$ any estimator with $s>1$ is dominated by the MLE $\hat{\mu}=X$

case 2 : $s<0$

$$\text{Risk}_k(|s|X) = (|s|-1)^2\|\mu\|^2 + |s|^2 d$$

$|s|>s$ when $s<0 \Rightarrow$ risk is always lower for $|s|X$ compared to sX

$\Rightarrow$ any estimator with $s<0$ is dominated by its positive counterpart $|s|X$. Correct +

(b) derivative with respect to s:

$$\frac{d}{ds}\left[(s-1)^2\|\mu\|^2+s^2 d\right] = 2(s-1)\|\mu\|^2+2sd$$

$$2(s-1)\|\mu\|^2+2sd=0$$

$$(s-1)\|\mu\|^2+sd=0$$

$$s\|\mu\|^2-\|\mu\|^2+sd=0$$

$$s(d+\|\mu\|^2)=\|\mu\|^2$$

$$s^* = \frac{\|\mu\|^2}{d+\|\mu\|^2} = 1-\frac{d}{d+\|\mu\|^2}$$

Yes +

(c) an estimator has to be a function of the observed data X but s^* depends on $\|\mu\|^2$ which is an unknown quantity.

modified estimator: $\left(1 - \frac{d}{\|X\|^2}\right) X$

$1 - \frac{d}{\|X\|^2}$ $\Rightarrow$ ensures that the shrinkage effect depends on magnitude of X

$\Rightarrow$ when $\|X\|^2$ is large, the shrinkage factor is close to 1

when $\|X\|^2$ is small, the shrinkage factor is significantly less than 1, pulling the estimator towards the origin.

+

(I would simply say that $\|X\|^2 - d$ is an unbiased estimate of $\|\mu\|^2$.)

(d) $R(\delta) = E_\mu \|\hat{\mu}^\delta - \mu\|^2$

$$\hat{\mu}^\delta - \mu = \left(1 - \frac{\delta}{\|X\|^2}\right)X - \mu$$

since $X = \mu + z$ and $z \sim N(0, I_d)$

$$\hat{\mu}^\delta - \mu = \left(1 - \frac{\delta}{\|X\|^2}\right)(\mu + z) - \mu$$

$$\hat{\mu}^\delta - \mu = \left(1 - \frac{\delta}{\|X\|^2}\right)\mu + \left(1 - \frac{\delta}{\|X\|^2}\right)z - \mu$$

$$= \left(1 - \frac{\delta}{\|X\|^2} - 1\right)\mu + \left(1 - \frac{\delta}{\|X\|^2}\right)z$$

$$= \left(-\frac{\delta}{\|X\|^2}\right)\mu + \left(1 - \frac{\delta}{\|X\|^2}\right)z$$

$$R(\delta) = E\left[\left\| -\frac{\delta}{\|X\|^2}\mu + \left(1 - \frac{\delta}{\|X\|^2}\right)z\right\|^2\right]$$

$$= E\left[\left\| -\frac{\delta}{\|X\|^2}\mu\right\|^2 + \left\|\left(1 - \frac{\delta}{\|X\|^2}\right)z\right\|^2 + \underbrace{2\left\langle -\frac{\delta}{\|X\|^2}\mu, \left(1 - \frac{\delta}{\|X\|^2}\right)z\right\rangle}_{=0}\right]$$

since, $E\langle\mu, z\rangle = 0$

$$\Rightarrow R(\delta) = E\left[\left\| -\frac{\delta}{\|X\|^2}\mu\right\|^2 + \left\|\left(1 - \frac{\delta}{\|X\|^2}\right)z\right\|^2\right]$$

$$= E\left[\frac{\delta^2}{\|X\|^4}\|\mu\|^2 + \left(1 - \frac{\delta}{\|X\|^2}\right)^2\|z\|^2\right]$$

since $E\|Z\|^2 = d$

$$R(\delta) = E\left[\frac{\delta^2}{\|X\|^4}\|\mu\|^2 + \left(1 - \frac{\delta}{\|X\|^2}\right)^2 d\right]$$

we want to minimize the quadratic risk

Stein lemma

$$\Rightarrow E[(X_i - \mu_i)g(X)] = E\left[\frac{\partial g}{\partial x_i}(X)\right]$$

$$\Rightarrow E\left[\frac{\delta}{\|X\|^2}\|\mu\|^2\right] = E\left[\frac{\delta d + \|\mu\|^2}{\|X\|^2}\right]$$

$$\frac{d}{d\delta}R(\delta) = E\left[\frac{2\delta}{\|X\|^4}\|\mu\|^2 - 2\left(1 - \frac{\delta}{\|X\|^2}\right)\frac{d}{\|X\|^2}\right]$$

$$= 0$$

$$\Rightarrow E\left[\frac{2\delta}{\|X\|^4}\|\mu\|^2\right] = E\left[\frac{2d}{\|X\|^2} - \frac{2\delta d}{\|X\|^4}\right]$$

$$E\left[\frac{2\delta}{\|X\|^4}\|\mu\|^2 + \frac{2\delta d}{\|X\|^4}\right] = E\left[\frac{2d}{\|X\|^2}\right]$$

$$E\left[\frac{2\delta(d + \|\mu\|^2)}{\|X\|^4}\right] = E\left[\frac{2d}{\|X\|^2}\right]$$

$$\delta^* = d - 2$$

plugging δ^* back into estimator

$$\Rightarrow \hat{\mu}^{JS} = \left(1 - \frac{d-2}{\|X\|^2}\right)X$$

+

Well done!

Problem 7

suppose we already have a valid planar Venn diagram for $n-1$ sets with F_{n-1} faces.
when adding the nth set, it must split every existing region it intersects into two.
since the previous Venn diagram had $F_{n-1} = 2^{n-1}$ face, adding a new set doubles the number of faces $\Rightarrow F_n = 2F_{n-1} \Rightarrow F_n = 2^n$

Also, each new curve adds at most $2(n-1)$ new regions
from Euler, for a planar graph:

$$V - E + F = 2$$

$$V_n - E_n \leq V_{n-1} - E_{n-1}$$

$$F_n < 2F_{n-1}$$

$$F_n = 2 + E_n - V_n$$

$$F_{n-1} = 2 + E_{n-1} - V_{n-1}$$

$$F_n \leq 2 + E_{n-1} - V_{n-1} + 2(n-1)$$

$F_n \leq F_{n-1} + 2(n-1)$ → recursive bound

iterating this gives: $F_n \leq 2\sum_{k=1}^{n-1} k = \frac{2(n-1)n}{2} = O(n^2)$

$$F_n \leq F_{n-1} + 2(n-1)$$

$$F_{n-1} \leq F_{n-2} + 2(n-2)$$

$$F_{n-2} \leq F_{n-3} + 2(n-3)$$

$$- \; - \; -$$

each new curve intersects the previous ones at most $2(n-1)$ times.

$\Rightarrow \boxed{V_n \leq V_{n-1} + 2(n-1)}$ +

for E_n:

each new curve must intersect all previous $n-1$ curves.

a simple closed curves intersects another simple closed curve at at least 2 points.

It creates $2(n-1)$ new edges.

However, if the Venn diagram is fully connected, each intersection divides the plane into more regions $\Rightarrow$ at least 4 edges per previous sets are needed.

$\Rightarrow \boxed{E_n \geq E_{n-1} + 4(n-1)}$ + Well done

---

for $n=3$, $F_3 = 2^3 = 8$

for $n=4$, $F_4 = 2^4 = 16$

However, $F_4 \leq F_3 + 2(4-1) \leq 8+6$
≤ 14

$\Rightarrow$ therefore you can't draw a Venn diagram for $n \geq 5$ by shifting a circle.

+