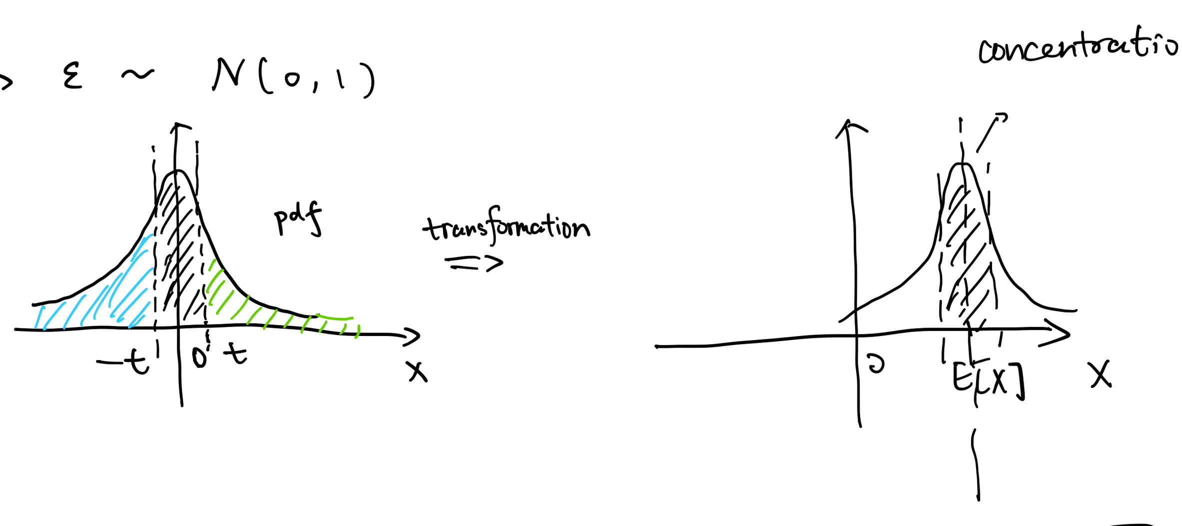
Week 5 Concentration

Thursday, February 8, 2024 1:55 PM

$X = E[X] + \varepsilon$

Ex. $X \sim$ Gaussian $N(0,1)$

$\Rightarrow \varepsilon \sim N(0,1)$

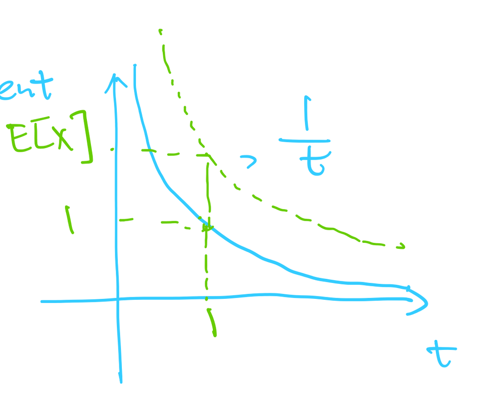


Concentration inequality : $\mathbb{P}(|X - E[X]| > t) \le \boxed{?}$ $(t>0)$

$\boxed{\mathbb{P}(X - E[X] > t)}$ & $\boxed{\mathbb{P}(X - E[X] < -t)}$

Markov inequality : $X \geq 0$

$\mathbb{P}(X > t) \le \dfrac{E[X]}{t}$ 1-moment



Ex : $X \in (-\infty, \infty)$ $X: \Omega \to \mathbb{R}$

$\mathbb{P}(\boxed{X > t}) = \mathbb{P}(\{w \in \Omega \mid X(w) > t\})$

$\Updownarrow$

$= \mathbb{P}(\{w \in \Omega \mid e^{sX(w)} > e^{st}\})$ $(s>0)$

$= \mathbb{P}(e^{sX} > e^{st})$

$\le \dfrac{E[e^{sX}]}{e^{st}}$ $= M_X(s)$ Bernstein's ineq.

$X = X\,\mathbb{1}\{X \geq 0\} + X\,\mathbb{1}\{X < 0\}$

Gaussian : $X \sim N(\mu, \sigma^2)$

pdf: $f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \boxed{e^{-\frac{1}{2\sigma^2}(x-\mu)^2}}$

mgf: k-moment : $E[X^k]$ moment generating function

$M_X(t) = E[e^{tX}] = E\left[\sum_{k=0}^{\infty} \dfrac{(tX)^k}{k!}\right] = \sum_{k=0}^{\infty} \dfrac{E[X^k]}{k!} t^k$

mgf: $M_X(t) = E[e^{tX}] = e^{\mu t + \frac{\sigma^2}{2}t^2}$

Markov inequality : $\mathbb{P}(X > t) \le \dfrac{E[e^{sX}]}{e^{st}}$

$= \dfrac{e^{\mu s + \frac{\sigma^2}{2}s^2}}{e^{st}}$

$f(s) = (\mu - t)s + \frac{\sigma^2}{2}s^2$

$f'(s) = \mu - t + \sigma^2 s = 0$

$\Rightarrow s^* = \dfrac{t-\mu}{\sigma^2}$

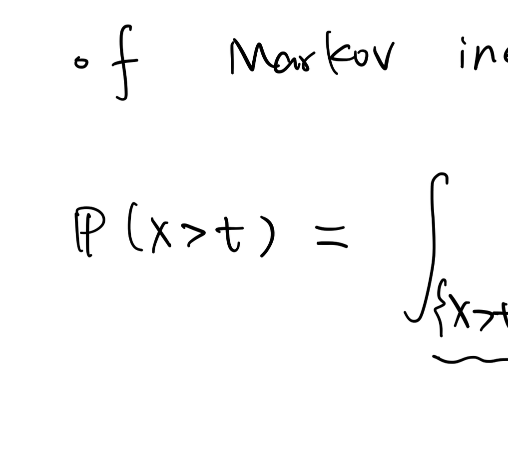
$= e^{(\mu-t)s + \frac{\sigma^2}{2}s^2}$ $\forall s > 0$

$\mathbb{P}(X > t) \le e^{(\mu-t)s + \frac{\sigma^2}{2}s^2}$ $\forall s > 0$

Let $s = s^* = \dfrac{t-\mu}{\sigma^2}$

$\Rightarrow \mathbb{P}(X > t) \le e^{-\frac{1}{\sigma^2}(t-\mu)^2 + \frac{\cancel{\sigma^2}}{2} \cdot \frac{(t-\mu)^2}{\cancel{\sigma^4}\ \sigma^2}}$

$= e^{-\frac{1}{2\sigma^2}(t-\mu)^2}$



Pf of Markov ineq. Key : $\mathbb{P}(X > t) = \mathbb{E}[\mathbb{1}\{X > t\}]$

$\mathbb{P}(X > t) = \displaystyle\int_{\{X>t\}} 1 \cdot dF(x) = \int \mathbb{1}\{X > t\}\, dF(x)$

$= E[\mathbb{1}\{X > t\}]$

$\le E\left[\dfrac{X}{t} \mathbb{1}\{X > t\}\right]$ $X \geq 0$

$\le E\left[\dfrac{X}{t}\right].$

Key Lemma: $X \sim N(\mu, \sigma^2)$

$\mathbb{P}(X > t) \le e^{-\frac{1}{2\sigma^2}(t-\mu)^2}$

$\Rightarrow \mathbb{P}(|X| > t) \le \mathbb{P}(\{X > t\} \cup \{X < -t\})$

$\le \mathbb{P}(X > t) + \mathbb{P}(X < -t)$

$\le e^{-\frac{1}{2\sigma^2}(t-\mu)^2} + e^{-\frac{1}{2\sigma^2}(-t-\mu)^2}$

when $\mu = 0$, $\mathbb{P}(|X| > t) \le 2e^{-\frac{t^2}{2\sigma^2}}$.

When $\lambda_n \geq 2\left\|\dfrac{X^T w}{n}\right\|_\infty$ θ^*: sparsity

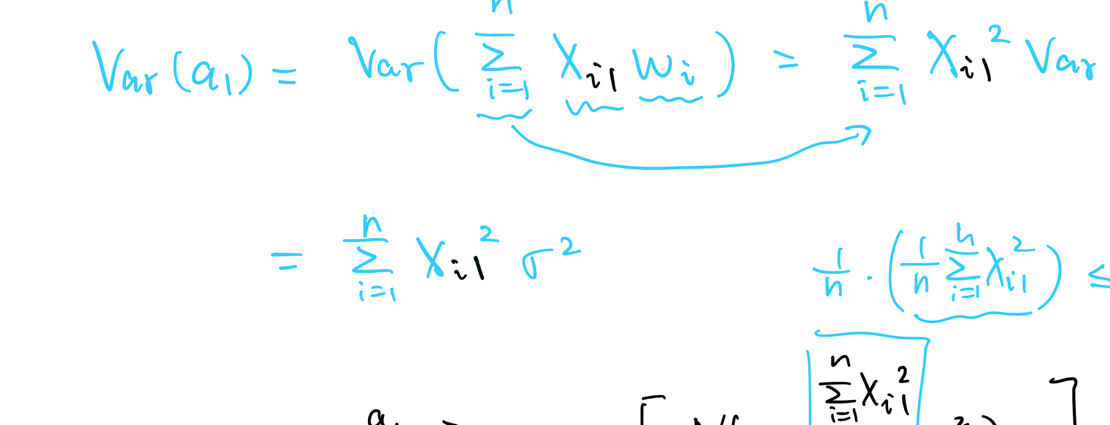
$\Rightarrow \boxed{\|\hat{\theta} - \theta^*\|_2^2 \le \dfrac{3}{\kappa}\sqrt{s}\,\lambda_n}$

RE - condition

$Y = X\theta^* + w$: if $w_i \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$, $w = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$

$\max\limits_{j=1,\dots,n} \dfrac{\|X(\cdot, j)\|_2}{\sqrt{n}} = \max\limits_{j=1,\dots,n} \dfrac{\sqrt{\sum_{i=1}^{n} X_{i,j}^2}}{\sqrt{n}} \le C.$

$$\begin{array}{cc} & d \\ X = n & \begin{bmatrix} X_{11} & \cdots & X_{1d} \\ X_{21} & \cdots & X_{2d} \\ \vdots & & \vdots \\ X_{n1} & \cdots & X_{nd} \end{bmatrix} \end{array}$$



$a_1 = \sum_{i=1}^{n} X_{1i} w_i \sim N\left(0, \sum_{i=1}^{n} X_{1i}^2 \sigma^2\right)$

$Var(a_1) = Var\left(\sum_{i=1}^{n} X_{i1} w_i\right) = \sum_{i=1}^{n} X_{i1}^2 Var(w_i)$

$= \sum_{i=1}^{n} X_{i1}^2 \sigma^2$ $\frac{1}{n} \cdot \left(\frac{1}{n}\sum_{i=1}^{n} X_{i1}^2\right) \le \boxed{\frac{C^2}{n}}$

$\left(\dfrac{X^T w}{n}\right) = \begin{bmatrix} \frac{a_1}{n} \\ \vdots \\ \frac{a_d}{n} \end{bmatrix} \sim \begin{bmatrix} N\left(0, \boxed{\frac{\sum_{i=1}^{n} X_{i1}^2}{n^2}}\sigma^2\right) \\ \vdots \\ N\left(0, \frac{\sum_{i=1}^{n} X_{id}^2}{n^2}\sigma^2\right) \end{bmatrix}$

$\mathbb{P}\left(\left\|\dfrac{X^T w}{n}\right\|_\infty > t\right) = \mathbb{P}\left(\max\left\{\left|\dfrac{X^T w(i)}{n}\right|\right\} > t\right)$

$\le \sum_{i=1}^{d} \mathbb{P}\left(\left|\dfrac{X^T w(i)}{n}\right| > t\right)$

$\le d \cdot 2e^{-\frac{t^2}{2C^2\sigma^2/n}}$

$= 2de^{-\boxed{\frac{nt^2}{2C^2\sigma^2}}} = \delta$

$\Rightarrow t = \sqrt{\log\left(\dfrac{2d}{\delta}\right) \cdot \dfrac{2C^2\sigma^2}{n}}$

$\le \boxed{\text{lec. Note}}$

with prob. at least $1 - \delta$

$\left\|\dfrac{X^T w}{n}\right\|_\infty \le \sqrt{\log\left(\dfrac{2d}{\delta}\right) \cdot \dfrac{2C^2\sigma^2}{n}}$

$\lambda_n \le 2\left\|\frac{X^T w}{n}\right\|_\infty$

$\boxed{\|\hat{\theta} - \theta^*\|_2 \le \dfrac{3}{\kappa}\sqrt{s}\,\lambda_n \le \dfrac{6}{\kappa} \cdot \sqrt{\log\left(\dfrac{2d}{\delta}\right) \cdot \dfrac{2C^2\sigma^2}{n}}}$