18.02.2020

Cramer – Chernoff Theory

$\lambda > 0, \quad P[X \geq t] = P[e^{\lambda X} \geq e^{\lambda t}] \leq \inf_{\lambda \geq 0} \frac{E[e^{\lambda X}]}{e^{\lambda t}}$

$P[X \leq t] \leq e^{-\psi_X^*(t)}$

$\ln P[X \geq t] \leq \inf_{\lambda \geq 0} \ln E[e^{\lambda X}] - \lambda t$

$\equiv - \sup_{\lambda \geq 0} \left(\lambda t - \ln E[e^{\lambda X}] \right) = \psi_X^*(t)$

Cramér – Chernoff Transformation

$\left(\psi_X^*(t) \right) \equiv \sup_{\lambda \geq 0} \left(t\lambda - \psi_X(\lambda) \right) = \psi_X^*(t)$

$\downarrow$ when $t \geq E[X]$

$\sup_{\lambda \in \mathbb{R}} \left(\lambda t - \psi_X(\lambda) \right)$ w

Fenchel Leptude Transformation

$f_X(x) \equiv \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \quad N(0, \sigma^2)$

$\psi_X(\lambda) \equiv \ln M_X(\lambda) = \ln e^{\frac{1}{2}\sigma^2\lambda^2} = \frac{1}{2}\sigma^2\lambda^2$

$\psi_X^*(t) = \sup_{\lambda \in \mathbb{R}} \left(\lambda t - \frac{1}{2}\sigma^2\lambda^2 \right) \quad t > EX = 0$

$t - \frac{1}{\cancel{2}}\cancel{2}\,\sigma^2 \cdot \lambda \Rightarrow t = \lambda\sigma^2$

$\lambda^* = \frac{t}{\sigma^2}$

$\psi_X^*(t) = \frac{t}{\sigma^2} \cdot t - \frac{1}{2} \cdot \sigma^2 \left(\frac{t}{\sigma^2} \right)^2 = \frac{t^2}{2\sigma^2}$

$$P[|X| \geq t] \leq 2e^{-\frac{t^2}{2\sigma^2}}$$

Definition: A random variable X is "sub-gamma" with variance σ^2

if $\Psi_X(t) \leq \frac{1}{2} t^2 \sigma^2$ $\quad M_X(t) \leq e^{\frac{1}{2}t^2\sigma^2}$.

Thm: Let X be a centered r.v (i.e, $E[X]=0$) and

$$P[X > x] \vee P[-X > x] \leq e^{-\frac{x^2}{2v}} \quad \text{for some } v > 0.$$

This result gives moment bound estimates of these kinds of random variables (in particular sub-gaussian)

Then $\forall$ integers $q \geq 1$, $E[X^{2q}] \leq 2 \times q!\,(2v)^q \overset{(*)}{\leq} q!\,(4v)^q$

$X \leftarrow \frac{X}{\sqrt{v}}$ (Scaling) ?

$$E[X^{2q}] = \int_0^\infty P[|X|^{2q} \overset{?}{>} x]\,dx$$

$$= 2q \int_0^\infty x^{2q-1} P[|X| > x]\,dx$$

$$\leq 4q \int_0^\infty x^{2q-1} e^{-x^2/2}$$

$x = \sqrt{2t}$

$$\leq 4q \int_0^\infty (2t)^{q-1} e^{t}\,dt$$

$$= \underset{?}{4}, 2^{q-1} \int_0^\infty t^{q-1} e^{-t}\,dt = 2^{q+1} \cdot q!$$

Claim: $E[X] = \sum_x P[X \geq x]$

$$P(A) = \int_A dP = \int_X I_A(\cdot)P(\cdot)$$

$$= E[I_A(\cdot)]$$

Then, $\sum_x P[X \geq x] = \sum_x E[I(X \geq x)]$

$$= E[\underbrace{\sum_x I(X \geq x)}_{X}] = E[X]$$

what happens if $(X < x)$?

$$\frac{2q!}{q!} = \prod_{j=1}^{q} (q+j) \geq \prod_{j=1}^{q} 2j = 2^q q!$$

$X \sim X'$ (copy of them)

$Y = X - X'$

What is the moment generating function of Y? $E[e^{\lambda(Y)}] = ?$

Claim: $\lambda > 0$, X's centered r.v. $E[e^{-\lambda X}] \geq 1$.

Proof: Jensen inequality will be used.

Jensen inequality:

$\underset{\text{convex}}{\phi}: \mathbb{R}^n \supseteq \underset{\text{convex}}{G} \longrightarrow \mathbb{R}$

$X = (X_1, X_2, \dots X_n) \in G$.

$E[\phi(X)] \geq \phi(E[X]) \longrightarrow E[\phi(X)] \geq \phi(E[X_1], \dots, E[X_n])$

Generalized Jensen

$\downarrow$

Minskovy

$\downarrow$

Cauchy Schwarz

- Real line is convex

$$E[e^{\lambda(-X)}] \geq e^{\lambda E[-X]} = 1$$

end of the proof

$Y = X - X'$

$$E[e^{tY}] = E[e^{tX}] \cdot \underbrace{E[e^{-tX}]}_{\geq 1}$$

$$\Rightarrow E[e^{tY}] \geq E[e^{tX}]$$

$$E[Y^{2q}] = E[(X-X')^{2q}]$$

— odd-powers are 0.

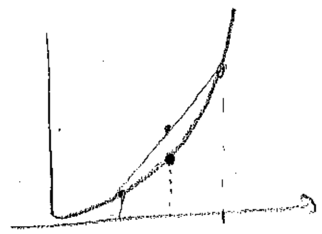
$$E[Y^{2q+1}] = E[(X-X')(X-X')^{2q}] =$$

$$\leq E[(|X|+|X'|)^{2q}] \leq 2^{2q} E[|X|^{2q}]$$

$$(|X|+|X'|)^{2q} = 2^{2q}\left(\tfrac{1}{2}|X| + \tfrac{1}{2}|X'|\right)^{2q} \leq 2^{2q}\left(\tfrac{1}{2}|X|^{2q} + \tfrac{1}{2}|X'|^{2q}\right)$$

not correct !!!
!!!

$x \longrightarrow x^{2q}$ convex



$$E[Y^{2q}] \leq 2^{2q} E[X^{2q}]$$

$$E[e^{tX}] \leq E[e^{tY}] \leq \sum_{q=0}^{n} \frac{2^{2q} E[X^{2q}]}{(2q)!} \cdots = \sum_{q=0}^{\infty} \frac{2^q q!}{(2q)!}\left(\frac{t^2(16v)}{2}\right)^q$$

$$\frac{2^q q!}{(2q)!} = \frac{2q\,(2q-2)(2q-4)\cdots \quad q!}{(2q)(2q-1)(2q-2)\cdots(q+1)(q\cdots 2.1)} \leq \frac{1}{q!}$$

$$\leq \sum_{q=0} \left(\frac{t^2(16v)}{2}\right)^q \frac{1}{q!} = e^{\frac{t^2(16v)}{2}} = e^{t^2(8v)} = e^{2t^2 \underline{4v}}$$

$\downarrow$
c

<u>Result</u>: If for some constant $c > 0$,

$E[X^{2q}] \leq q!\, c^q$ then $X \sim$ subgaussian with $v = 4c$.

1. $E[X^{2q}] \leq q!(4v)^q$.

2. If $c > 0$, $E[X^{2q}] \leq q!\,c^q$

   └ subgaussian $c = 4v$

- the ~~sec.~~ n-th order moment estimates : $E[X^n] \sim \left(\frac{n}{2}\right)!$. vary / varies.

Johnson - Lindenstrauss Theorem : (query optimization. - you're searching smt, then it goes to database and presents the most relevant)

- recommendations
- search engines.
  (nearest neighbour)
  It allows to keep distances in original way in projection space.

- Take a suitable r.v.
  random ↓ Projection.
- gaussian & sub-gaussian.

- the beauty: the projection depends only number of data points not the dimension of vector space.

- Inf. dim space, we can also do projections : If we have a set, one way to measure is cardinality. Are there any other? Yes.

* We are still in finite domain.

  Let $A \subset R^d$ be finite set $|A| = N$. Assume $v \geq 1$.

  $X_{ij} \in \mathcal{G}(v)$. $\mathcal{G}(v)$ = subgaussian with var v.

  Let $\varepsilon, \delta \in (0,1)$. Then if $K \geq 100 N^2 \varepsilon^{-2} \ln\left(\frac{N}{\sqrt{\delta}}\right)$ then

  w.p $1 - \delta$ (at least) W is an ε-isometry on A.

$$S_d \supset T \equiv \left\{ \frac{v - v'}{\|v - v'\|} : v \neq v', \, v, v' \in A \right\}$$

– What's the cardinality of T? $\underbrace{|T| \leq \binom{n}{2}}_{\text{why?}}$

– $\sup_{v \in T} \left| \|\Pi(v)\|^2 - 1 \right| \leq \varepsilon$

ε-isometry

$v \in T \Rightarrow \|v\| = 1$

that's why not the $\|v\|$ but 1

$$\underbrace{E\left[e^{\lambda \Pi_i(v)}\right]}_{\text{dict projections}} = E\left[e^{\lambda \sum_{j=1}^{d} v_j X_{ij}}\right] = \prod_{j=1}^{d} E\left[e^{\lambda v_j \overbrace{X_{ij}}^{\text{subgaussian}}}\right]$$

$$\leq e^{\lambda^2 \sum_{j=1}^{d} v_j^2 v/2}$$

$$= e^{\lambda^2 v/2}$$

dict projections ? are also sub-gaussians.

Therefore, $\Pi_i(v) \in \overbrace{\mathcal{G}(v)}^{\text{subgaussian}} \Rightarrow E\left[\Pi_i(v)^{2q}\right] \leq \frac{q!}{2}(4v)^2$

Pf of the first generalization of Johnson-Lindenstrauss Thm.