High Dimension Probability - On Concentration Inequalities

Bipin Koirala

March 2023

1 Gaussian Concentration Inequality

Let, $X \sim \mathcal{N}(0, I_d) \in R^d$ and $f : R^d \to R$ a Lipschitz function i.e. $|f(X) - f(y)| \leq L \; ||x - y||$. Then;

$$\boxed{\mathbb{P}\big\{|f(X) - \mathbb{E}f(X)| \geq t\big\} \leq 2 \text{ exp } (-t^2/2L^2) \;\; ; \forall \; t > 0}$$

Remark

- $f(x) = \langle w, x \rangle$ then $L = ||w||$. If $X \sim \mathcal{N}(0, I_d)$ then $f(x) = \langle w, x \rangle \sim \mathcal{N}(0, ||w||^2)$ and

$$\mathbb{P}\{f(X) \geq t\} = \mathbb{P}\{f(X) - \mathbb{E}f(X) \geq t\} = 1 - \Phi\bigg(\frac{t}{||w||}\bigg) \text{ if Median or } \mathbb{E}f(X) = 0$$

  where; $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-y^2/2} \; dy$ and $1 - \Phi(t) \leq \frac{1}{2} e^{-t^2/2}$.

  $\to \mathbb{P}\{|f(X) - \text{Med } f(X)| \geq t\} \leq 2\bigg[1 - \Phi\big(\frac{t}{L}\big)\bigg] = \int_0^\infty e^{-t^2/2L^2} dt = \sqrt{\frac{\pi}{2}} L$

  $\to \mathbb{P}\{|f(X) - \text{Med } f(X)| \geq tL\} \leq 2\bigg[1 - \Phi\big(t\big)\bigg] \implies \mathbb{P}\big\{|f(X) - \mathbb{E}f(X)| \geq (t + \sqrt{\frac{\pi}{2}})L\big\} \leq 2(1 - \Phi(t))$

- If $X \sim \mathcal{N}(0, \Sigma)$ then $X = \Sigma^{1/2} Z$ where $Z \sim \mathcal{N}(0, I_d)$ then;

$$\mathbb{P}\{|f(X) - \mathbb{E}f(X)| \geq t\} \leq 2 \text{ exp}\bigg(-\frac{t^2}{2L^2 ||\Sigma||}\bigg)$$

2 Gaussian Concentration and Isoperimetry

Among all subsets of same surface area; sphere(ball) has the largest volume. Among all subsets of same surface volume; Euclidean ball has the smallest surface area.

Isoperimetry for Gaussian Measure $\gamma(A) : = \int_A \frac{1}{(2\pi)^d/2} e^{-|x|^2/2 \; dx}$; $A \subset \mathbb{R}^d$ is a Borel set.
i.e. γ is the standard Gaussian measure on Borel σ algebra in $\mathbb{R}^d$. γ is the distribution of $X \sim \mathcal{N}(0, I_d)$.

$$\boxed{\gamma(A) = \mathbb{P}(X \in A)}$$

Let; $A, H \subset \mathbb{R}^d$ be Borel subsets with $H = \{x \in \mathbb{R}^d : \langle w, x \rangle \leq c\}$ with $||w|| = 1$ and $c \in \mathbb{R}$ then;

$$\gamma(A) = \gamma(H) \implies \gamma(A_\epsilon) \geq \gamma(H_\epsilon) \;\; ; \; \forall \; \epsilon > 0$$

$\gamma(H_\epsilon) = \gamma(H_{w,c+\epsilon}) = \Phi(c+\epsilon)$.

$$\boxed{\Phi^{-1}(\gamma(A_\epsilon)) \geq \Phi^{-1}(\gamma(A)) + \epsilon}$$

Suppose; $\gamma(A) \geq 1/2$. Then $\Phi^{-1}(\gamma(A)) \geq \Phi^{-1}(1/2) = 0$ *i.e.* $\Phi^{-1}(\gamma(A)) \geq 0$ so $\gamma(A\epsilon) \geq \Phi(\epsilon)$

$$\gamma((A_\epsilon)^c) = 1 - \gamma(A_\epsilon) \leq 1 - \Phi(\epsilon)$$

$$\boxed{\gamma((A_t)^c) = 1 - \gamma(A_t) \leq \frac{1}{2}e^{-t^2/2}}$$

3 Application of Gaussian Concentration

3.1 Gaussian Stochastic Processes

Note: We're interested in concentration of $\sup\limits_{t\in\tau} |X(t)| := ||X||_\tau$

Let; $X(t)$ s.t. $t \in \tau$ be a centered Gaussian Process. Suppose that $\sup_{(t\in\tau)} |X(t)| < +\infty$. Let M be a median of R.V. $\sup_{t\in\tau} |X(t)|$. Denote: $\sigma^2 := \sup_{t\in\tau} \mathbb{E}\big[X^2(t)\big]$. Then following are true:

i $\mathbb{P}\{\sup_{t\in\tau} |X(t)| \geq M + u\} \leq 1 - \Phi(u/\sigma)$

ii $\mathbb{P}\{\sup_{t\in\tau} |X(t)| \leq M - u\} \leq 1 - \Phi(u/\sigma)$

iii $\mathbb{P}\{\big|\sup_{t\in\tau} |X(t)| - M\big| \geq u\} \leq 2\big(1 - \Phi(u/\sigma)\big); \quad \forall\ u \geq 0$

Corollary:

If; $\left\{\big|\sup_{(t\in\tau)} |X(t)| - M\big| \geq u\right\} \leq e^{-u^2/(2\sigma^2)}$
$\implies \mathbb{E}\big|\sup_{(t\in\tau)} |X(t)| - M\big| \leq \int_0^\infty e^{-u^2/(2\sigma^2)}\ du = \sqrt{\frac{\pi}{2}}\sigma$

$\implies \mathbb{E}\sup_{t\in\tau} |X(t)| \leq M + \sqrt{\frac{\pi}{2}}\sigma \leq +\infty$

Furthermore; $\mathbb{E}\ \exp\ \{\lambda \sup_{t\in\tau} |X(t)|^2\} < \infty; \quad \lambda < \frac{1}{2\sigma^2}$

3.2 Concentration Inequalities for norms of Gaussian R.V in Banach Spaces

E is a Banach Space and E^* is its dual space which consists of all bounded linear functionals on E. A bounded linear functional is a linear transformation that does not "blow up" the size of vectors in the vector space it is defined on. $||T(x)|| \leq M||x||$.

Let; $U \in E^* \Leftrightarrow X \in E \to \langle x, u\rangle \in \mathbb{R}$.

$$||U|| = \sup_{||x||\leq 1} \langle x, u\rangle < +\infty$$

Suppose, $M \subset \{U \in E^* : ||U|| \leq 1\}$. X is Gaussian $\sim \mathcal{N}(\mu, \Sigma) \Leftrightarrow \forall\ U \in E^*$; $\langle x, u\rangle$ is a normal R.V. (stochastic). Σ is covariance operator of X and

$$\langle \Sigma u, v\rangle = \text{Cov}\ \big(\langle x, u\rangle, \langle x, u\rangle\big)\ \ ;\ U, V \in E^*$$

We ar interested in concentration inequalities for

$$||X|| = \sup_{U \in M} |\langle x, u \rangle|$$

Assume that X is centered and $\sigma^2 := \sup_{U \in M} \mathbb{E}\langle x, u\rangle^2 \leq ||\Sigma||$. Let M be the median of X. Then following are true;

i $\mathbb{P}\{||X|| \geq M + U\} \leq 1 - \Phi(U/\sigma)$

ii $\mathbb{P}\{||X|| \leq M - U\} \leq 1 - \Phi(U/\sigma)$

iii $\mathbb{P}\{\big|||X|| - M\big| > U\} \leq 2\big(1 - \Phi(U/\sigma)\big)$; $\forall\, U \geq 0$

For all centered Gaussian R.V. $X \in E$;

$$\mathbb{E}||X|| < +\infty$$

$$\mathbb{E} \exp \{\lambda ||X||^2\} < \infty \quad ; \lambda < 1/(2\sigma^2)$$

3.3 Concentration of Lipschitz functions on the Sphere in $\mathbb{R}^n$

Let; $X \sim \mathrm{Unif}(\sqrt{n}S^{n-1})$ and Lipshitz $f : \sqrt{n}S^{n-1} \to \mathbb{R}$. Then, $||f(X) - \mathbb{E}f(X)||_{\psi_2} \leq \frac{||f||_{Lip}}{\sqrt{n}}$ Furthermore, $\mathbb{P}\{|f(X) - \mathbb{E}f(X)| \geq t\} \leq 2 \exp\big(-\frac{ct^2}{||f||^2_{Lip}}\big)$

4 Johnson-Lindenstrauss Lemma

Let $F = \{x_1, ..., x_n\}$ be a finite set of points in $\mathbb{R}^n$. With probability at least $1 - 2\exp(-c\epsilon^2 m)$, a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ is an ϵ-isometry $\big(\epsilon \in (0,1)\big)$ from F into $\mathbb{R}^m$ iff

$$\boxed{(1-\epsilon)||x-y||_2 \leq ||Ax - Ay||_2 \leq (1+\epsilon)||x-y||_2 \text{ for all } \; x, y \in F}$$

where $A = \sqrt{\frac{n}{m}} P_L$, P_L is an orthonormal projection onto a random m-dimensional subspace $L \subset \mathbb{R}^n$ and $m \geq (C\epsilon^2) \log N$.

5 Concentration of Functions of Independent R.V.

Suppose, $X_1, ..., X_n$ are independent R.V. with values in some spaces $S_1, ..., S_n$.
Let, $f : S_1 \times ..... \times S_n \to \mathbb{R}$ and let $Z := f(X_1, ...X_n)$. We're interested in the concentration of Z around its expectation (or median).