Gaussian Distributions (Recap)

"Standard Normal" $X \sim \mathcal{N}(0,1)$

density $\gamma(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

$\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = \mathrm{Var}[X] = 1$, $\mathbb{E}\left[e^{\lambda \cdot X^2}\right] = \frac{1}{\sqrt{1-2\lambda}}$

(for $\lambda < \frac{1}{2}$)

Multiple independent $X_1 \ldots\ldots X_n \sim \mathcal{N}(0,1)$

- $c_1 \cdot X_1 + c_2 \cdot X_2 + \ldots\ldots + c_n \cdot X_n + \mu \sim \mathcal{N}(\mu, \|c\|^2)$

- $\gamma(x_1 \ldots\ldots x_n) = \frac{1}{(2\pi)^{n/2}} e^{-(x_1^2 + \ldots + x_n^2)/2} = \frac{1}{(2\pi)^{n/2}} e^{-\|x\|^2/2}$

Johnson-Lindenstrauss dimension reduction [JL84]

- Let $P \subseteq \mathbb{R}^d$ be a set of points. $|P| = n$

$\forall \epsilon > 0$ $\exists$ linear transformation $\varphi : \mathbb{R}^d \longrightarrow \mathbb{R}^k$
$(\epsilon < 1)$

$$k \leq \frac{12 \ln n}{\epsilon^2 - \epsilon^3}$$

want $k \ll d$

s.t. $\forall u, v \in P$

$$(1-\epsilon) \cdot \|u - v\|^2 \leq \|\varphi(u) - \varphi(v)\|^2 \leq (1+\epsilon) \cdot \|u - v\|^2$$

Designing φ (at random)

- Pick (random) matrix $G \in \mathbb{R}^{k \times d}$ s.t.

  $\forall\, i \in [k],\ j \in [d].\quad G_{ij} \sim \mathcal{N}(0,1)$

  & all entries are independent

$$\begin{array}{cc} & \overbrace{\hphantom{-\ -\ -\ -}}^{d} \\ k\left\{\vphantom{\begin{bmatrix} - \\ \\ \\ \end{bmatrix}}\right. & \begin{bmatrix} - & - & - \\ & & \\ & & \end{bmatrix} \end{array}$$

- $\forall\, u \in \mathbb{R}^d$, define $\varphi(u) = \dfrac{Gu}{\sqrt{k}}$

- Fix $u, v \in P$. Suffices to prove

$$\mathbb{P}_G\left[(1-\epsilon)\|u-v\|^2 \leq \|\varphi(u)-\varphi(v)\|^2 \leq (1+\epsilon)\cdot\|u-v\|^2 \right] \geq 1 - \frac{2}{n^3}$$

for $k \geq \frac{12 \ln n}{\epsilon^2 - \epsilon^3}$

$$\mathbb{P}_G\left[\forall u,v \in P \quad (1-\epsilon)\|u-v\|^2 \leq \|\varphi(u)-\varphi(v)\|^2 \leq (1+\epsilon)\|u-v\|^2 \right]$$

$$\geq 1 - \binom{n}{2}\cdot\frac{2}{n^3}$$

$$\geq 1 - \frac{1}{n}$$

$$\mathbb{P}\left[(1-\epsilon)\|u-v\|^2 \le \|\varphi(u)-\varphi(v)\|^2 \le (1+\epsilon)\cdot\|u-v\|^2\right]$$

$$\parallel$$

$$\mathbb{P}\left[(1-\epsilon)\cdot\|u-v\|^2 \le \frac{\|G(u-v)\|^2}{k} \le (1+\epsilon)\cdot\|u-v\|^2\right]$$

$$\mathbb{P}\left[(1-\epsilon)\cdot k \le \frac{\|G(u-v)\|^2}{\|u-v\|^2} \le (1+\epsilon)\cdot k\right]$$

$$w = \frac{u-v}{\|u-v\|}$$

$$\mathbb{P}\left[(1-\epsilon)\cdot k \le \|Gw\|^2 \le (1+\epsilon)\cdot k\right]$$

Fix $w \in \mathbb{R}^d$, $\|w\| = 1$

$$k\left\{\overbrace{\begin{bmatrix} & & \\ & G & \\ & & \end{bmatrix}}^{d}\right. \begin{bmatrix} \\ w \\ \\ \end{bmatrix} = \begin{bmatrix} w_1 \cdot G_{11} + w_2 \cdot G_{12} + \cdots + w_d \cdot G_{1d} \\ \vdots \\ w_1 \cdot G_{i1} + \cdots + w_d \cdot G_{id} \end{bmatrix}$$

$$X_i = w_1 \cdot \overbrace{G_{i1}}^{\mathcal{N}(0,1)} + \cdots + w_d \cdot \overbrace{G_{id}}^{\mathcal{N}(0,1)} \sim \mathcal{N}(0, \underbrace{\|w\|^2}_{1})$$

$$\|Gw\|^2 = Z = X_1^2 + \cdots + X_k^2$$

Sums of independent squared Gaussians

$$Z = X_1^2 + \cdots + X_k^2$$

$X_1, \ldots, X_k \sim N(0,1)$

independent

$\mathbb{E}[Z] = k$

$$\boxed{\begin{gathered}\lambda < 1/2 \\ (\lambda > 0)\end{gathered}}$$

$$P[Z \geq (1+\epsilon)\cdot k] = P\left[e^{\lambda Z} \geq e^{\lambda(1+\epsilon)k}\right]$$

$$e^{\lambda Z} = \prod_{i=1}^{k} e^{\lambda X_i^2}$$

$$\leq \frac{\mathbb{E}[e^{\lambda Z}]}{e^{\lambda(1+\epsilon)k}}$$

$$= \frac{\prod_{i=1}^{k} \mathbb{E}[e^{\lambda X_i^2}]}{e^{\lambda(1+\epsilon)k}} = \frac{?}{e^{\lambda(1+\epsilon)k}}$$

$$\mathbb{P}[Z \geq (1+\epsilon)\cdot k] \leq \frac{\left(\frac{1}{\sqrt{1-2\lambda}}\right)^{k}}{e^{\lambda(1+\epsilon)k}} = \left(\frac{e^{-2\lambda(1+\epsilon)}}{1-2\lambda}\right)^{k/2}$$

choose $\lambda = \dfrac{\epsilon}{2(1+\epsilon)}$

$$\left(e^{-\epsilon}(1+\epsilon)\right)^{k/2} \leq \left(\left(1-\epsilon+\frac{\epsilon^2}{2}\right)(1+\epsilon)\right)^{k/2}$$

$$= \left(1-\epsilon^2+\frac{\epsilon^2}{2}+\frac{\epsilon^3}{2}\right)^{k/2}$$

$$\leq e^{-(\epsilon^2-\epsilon^3)k/4}$$

Ex: $\mathbb{P}[Z \leq (1-\epsilon)k] \leq e^{-(\epsilon^2-\epsilon^3)k/4}$

Putting things together

$$\mathbb{P}\left[\,\Big|\,\underbrace{\|Gv\|^2}_{Z} - k\,\Big| \geq \epsilon k\right] \leq 2 \cdot e^{-(\epsilon^2-\epsilon^3)\cdot k/4}$$

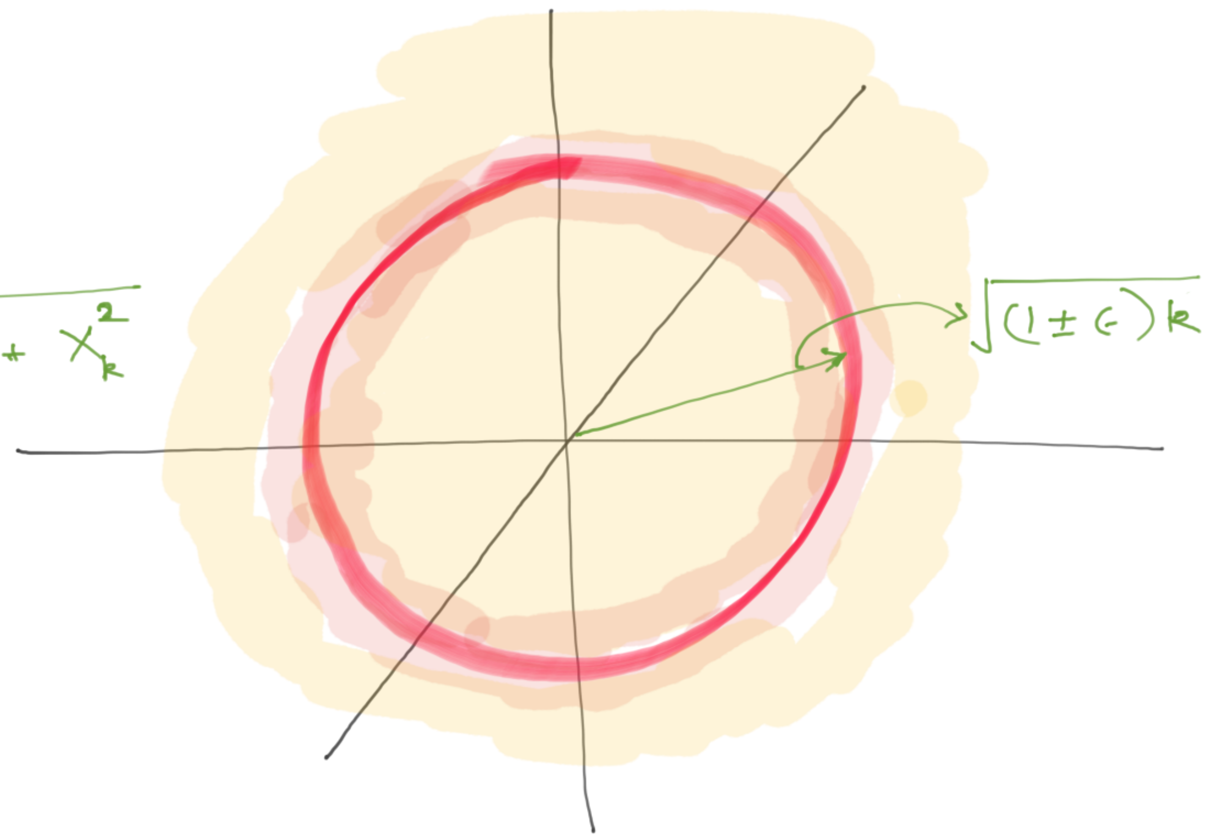
$$\leq \frac{2}{n^3} \qquad k \geq \frac{12 \ln n}{\epsilon^2 - \epsilon^3}$$

$$\mathbb{P}_{G}\left[\exists u, v \in \mathcal{P} \;.\; \left|\frac{\|G(u-v)\|^2}{k} - \|u-v\|^2\right| \geq \epsilon \|u-v\|^2\right] \leq$$

Multi-dimensional Spherical Gaussians

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$$

$$\|x\| = \sqrt{x_1^2 + \cdots + x_k^2}$$



$$\frac{e^{-\|x\|^2/2}}{(2\pi)^{k/2}}$$

Gordon's thm : infinite union bounds

$T \subseteq \mathbb{R}^d \cap \{w \mid \|w\| = 1\}$

$$\mathbb{P}\left[\exists w \in T, \left|\frac{\|Gw\|^2}{k} - 1\right| \geq \epsilon\right] \leq \ldots\ldots$$

$\gamma(T) \equiv$ Gaussian width

$$\mathop{\mathbb{E}}_{x \sim \mathcal{N}(0,I)} \sup_{w \in T} |\langle x, w \rangle|$$

Dealing with general Gaussian vectors

Let $X_1, \ldots\ldots, X_n \sim N(0,1)$ be independent.

$$\underline{Y} = A\underline{X} + \mu$$

Let

$$\underline{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} & & \\ & A & \\ & & \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$$

$$\underbrace{\mathbb{E}[\underline{Y}]}_{\text{mean (vector)}} = \mu$$

$$\underbrace{\mathbb{E}\left[(\underline{Y} - \mathbb{E}[\underline{Y}])(\underline{Y} - \mathbb{E}[\underline{Y}])^T\right]}_{\text{covariance (matrix)}} = AA^T$$

Dealing with general Gaussian vectors

Let $\underline{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$ be a Gaussian random vector

With $\mathbb{E}[\underline{Y}] = \mu$ $\qquad \mathbb{E}\left[(\underline{Y}-\mu)(\underline{Y}-\mu)^T\right] = \Sigma$

Then $\underline{Y}$ can expressed as $\underline{Y} = \Sigma^{1/2} \cdot X + \mu$

for $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \sim (\mathcal{N}(0,1))^n$.

density $\gamma(\underline{Y}) = \,?$

Diagonalizing Σ

Let $\Sigma = \sigma_1^2 \cdot \underbrace{w_1 \cdot w_1^*}_{} + \cdots\cdots + \sigma_n^2 \cdot \underbrace{w_n \cdot w_n^*}_{}$

orthonormal basis of eigenvectors

$$\Sigma^{1/2}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} | & & | \\ w_1 & \cdots & w_n \\ | & & | \end{pmatrix}\begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix}\begin{pmatrix} - & w_1^* & - \\ & \vdots & \\ - & w_n^* & - \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \begin{pmatrix} | & & | \\ w_1 & \cdots & w_n \\ | & & | \end{pmatrix}\begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix}\begin{pmatrix} \langle w_1, x\rangle \\ \vdots \\ \langle w_n, x\rangle \end{pmatrix} \begin{matrix} \to x_1' \\ \\ \to x_n' \end{matrix}$$

$$\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = \begin{pmatrix} — w_1^* — \\ \vdots \\ — w_n^* — \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = W^* \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$\underline{Y} = \begin{pmatrix} | & & | \\ w_1 & \cdots & w_n \\ | & & | \end{pmatrix} \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix} \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = W \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix} x'$$

$$\mathbb{E}[X'] = 0 \qquad \mathbb{E}[(X' - \mathbb{E}X')(X' - \mathbb{E}X')] = I$$

Multi-dimensional Gaussians