بسم الله الرحمن الرحیم

جلسه دوازدهم

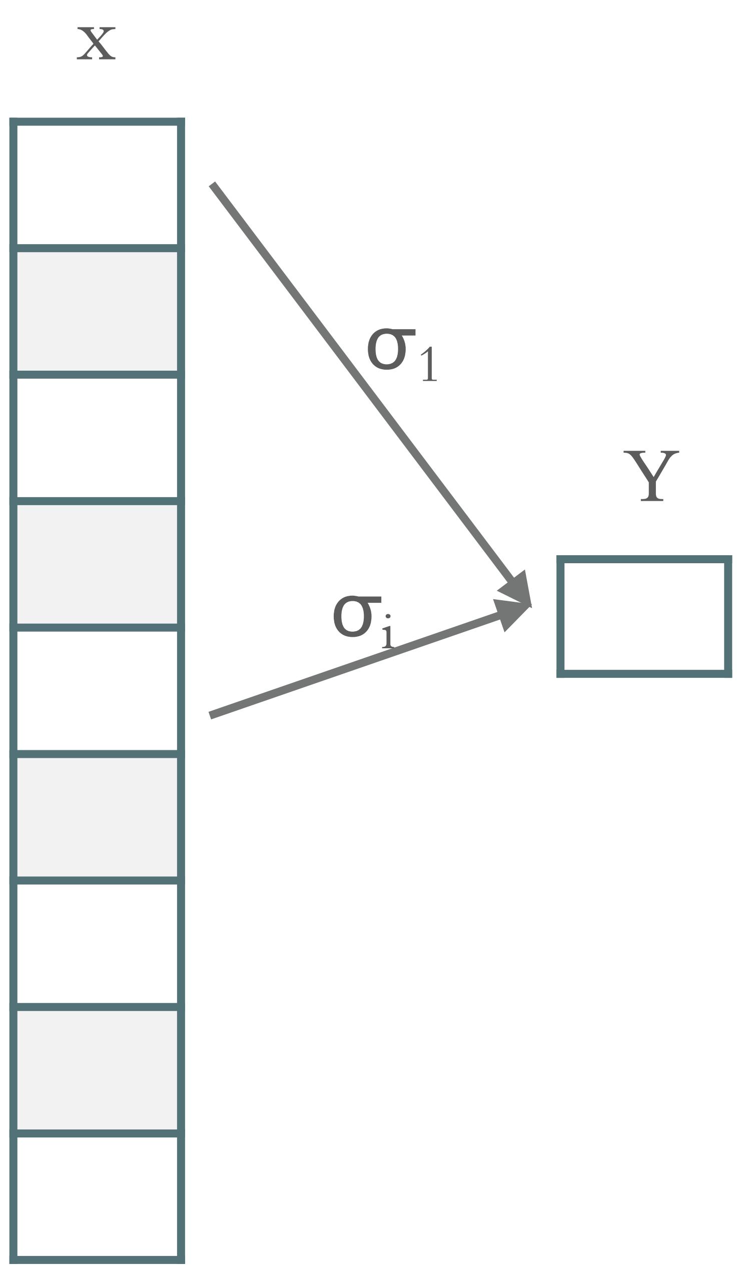
خلاصه‌سازی برای مه‌داده

مرور

Johnson–Lindenstrauss Transforms

ایده AMS برای تخمین نرم ۲:

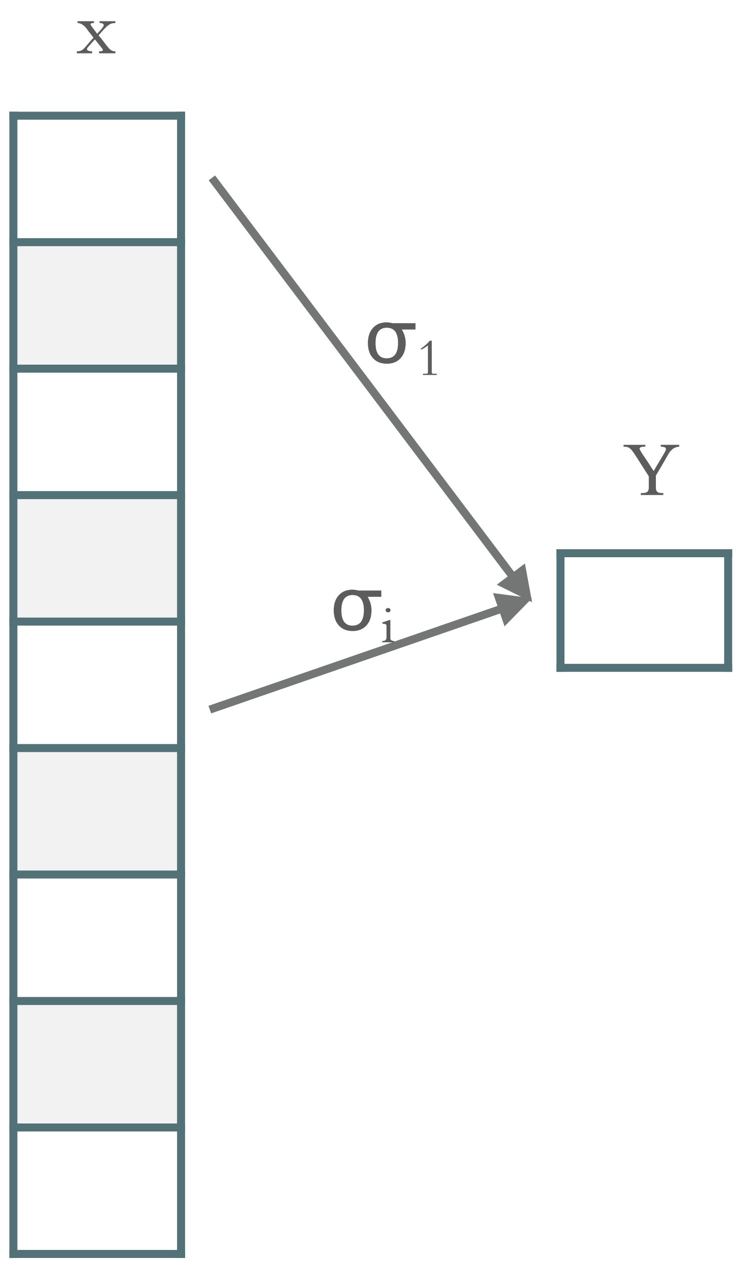
σ_i = +1 یا −1 (تصادفی یکنواخت)



ایده AMS برای تخمین نرم ۲:

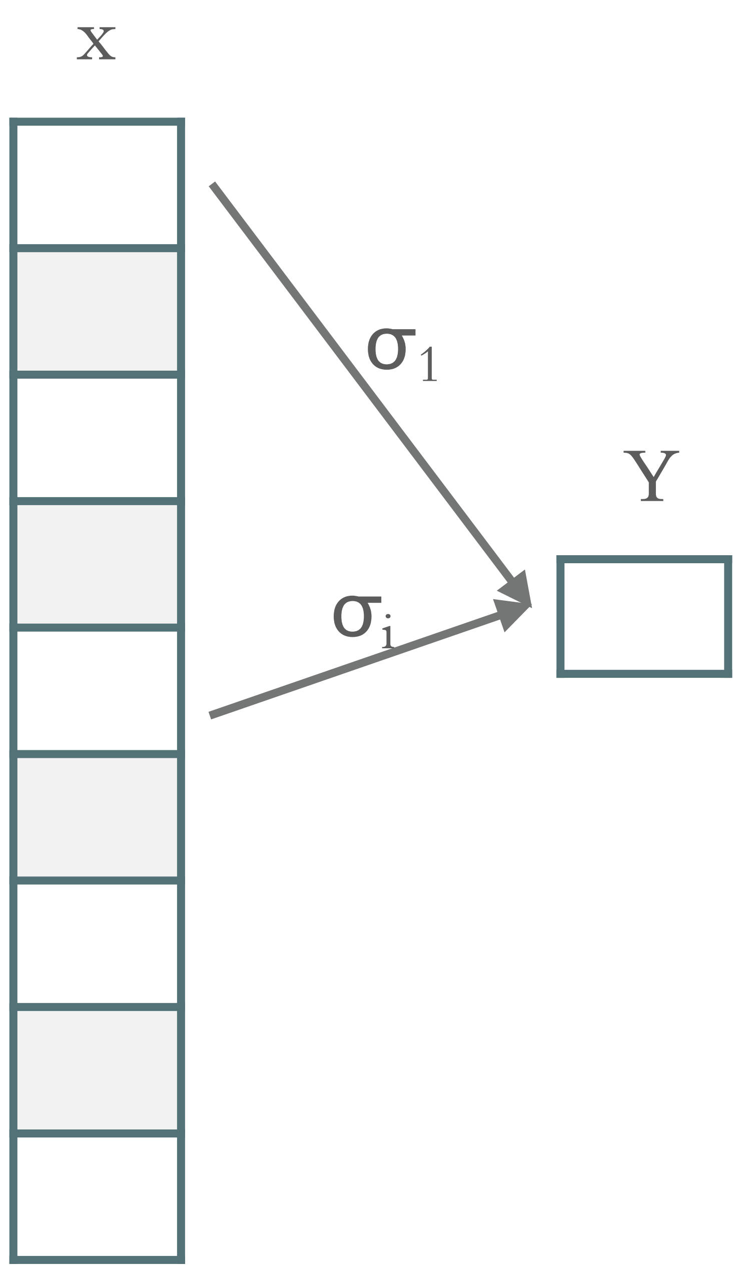
$\sigma_i = +1$ یا -1 (تصادفی یکنواخت)

$EY = 0$



ایده AMS برای تخمین نرم ۲:

$\sigma_i = +1$ یا -1 (تصادفی یکنواخت)



$$EY = 0$$

$$EY^2 = E\left[\sum_{j,j'} \sigma_j \sigma_{j'} x_j x_{j'}\right]$$

$$= E\left[\sum_{j} \sigma_j^2 x_j^2 + \sum_{j \neq j'} \sigma_j \sigma_{j'} x_j x_{j'}\right]$$

$$= \sum_{j} E[x_j^2] + \sum_{j \neq j'} E[\sigma_j \sigma_{j'} x_j x_{j'}]$$

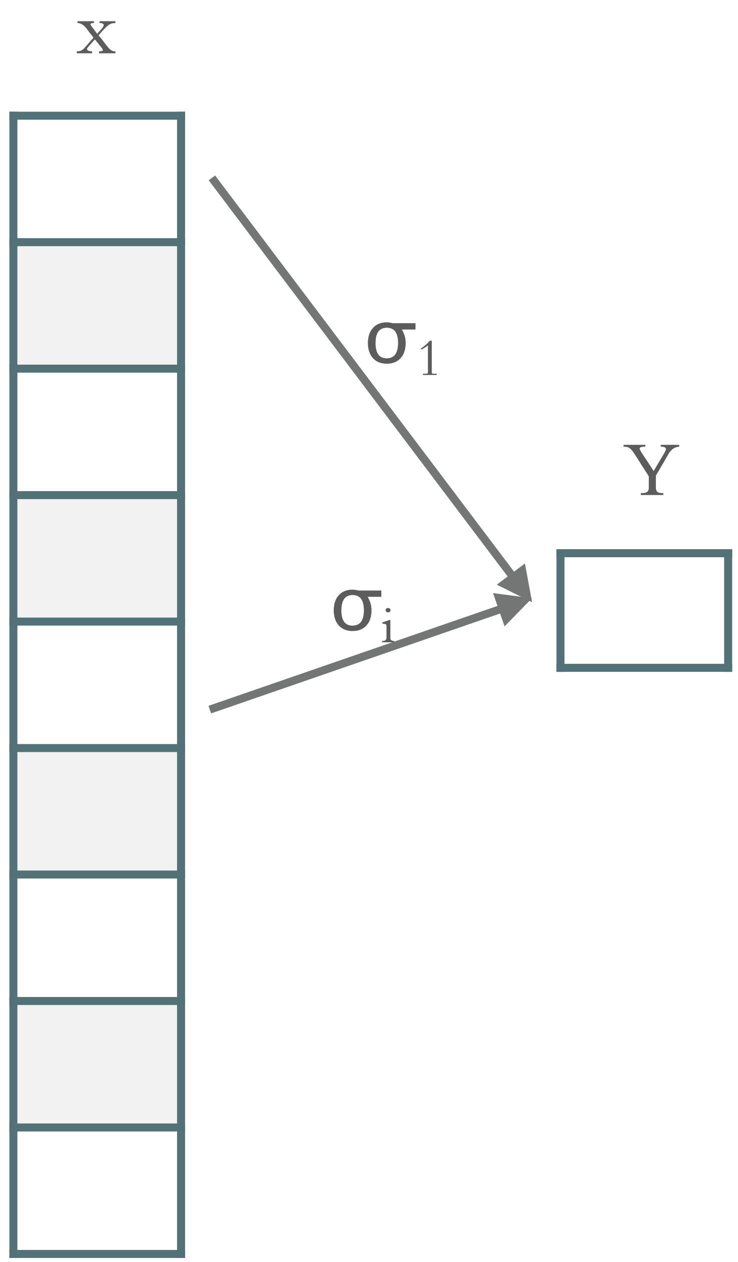
$$= ||x||_2^2$$

$E[\sigma_j]E[\sigma_{j'}]x_j x_{j'}$

و $E[\sigma_j] = 0$

ایده AMS برای تخمین نرم ۲:

$\sigma_i = +1$ یا -1 (تصادفی یکنواخت)



$$EY = 0$$

$$EY^2 = E\left[\sum_{j,j'} \sigma_j \sigma_{j'} x_j x_{j'}\right]$$
$$= E\left[\sum_{j} \sigma_j^2 x_j^2 + \sum_{j \neq j'} \sigma_j \sigma_{j'} x_j x_{j'}\right]$$
$$= \sum_{j} E[x_j^2] + \sum_{j \neq j'} E[\sigma_j \sigma_{j'} x_j x_{j'}]$$
$$= ||x||_2^2$$

$E[\sigma_j]E[\sigma_{j'}]x_j x_{j'}$

و $E[\sigma_j] = 0$

چرا برای بردار معمولی استفاده نکنیم؟

مرور

مرور

نسبتا خوب می‌شد: فاصله‌ها را تقریبا حفظ کند

$$\forall x, y \in X,\ \ (1-\varepsilon)\|x-y\|_2^2 \leq \|f(x)-f(y)\|_2^2 \leq (1+\varepsilon)\|x-y\|_2^2$$

مرور

نسبتا خوب می‌شد: فاصله‌ها را تقریبا حفظ کند

$$\forall x, y \in X,\ \ (1-\varepsilon)\|x-y\|_2^2 \le \|f(x)-f(y)\|_2^2 \le (1+\varepsilon)\|x-y\|_2^2$$

Theorem 5.0.1 (JL lemma [JL84]). *For any $\varepsilon \in (0,1)$ and any $X \subset \mathbb{R}^d$ for $|X| = n$ finite, there exists an embedding $f : X \to \mathbb{R}^m$ for $m = O(\varepsilon^{-2}\log n)$ such that*

$$\forall x, y \in X,\ \ (1-\varepsilon)\|x-y\|_2^2 \le \|f(x)-f(y)\|_2^2 \le (1+\varepsilon)\|x-y\|_2^2. \tag{5.1}$$

مرور

نسبتا خوب می‌شد: فاصله‌ها را تقریبا حفظ کند

$$\forall x, y \in X,\ (1-\varepsilon)\|x-y\|_2^2 \le \|f(x)-f(y)\|_2^2 \le (1+\varepsilon)\|x-y\|_2^2$$

Theorem 5.0.1 (JL lemma [JL84]). *For any $\varepsilon \in (0,1)$ and any $X \subset \mathbb{R}^d$ for $|X| = n$ finite, there exists an embedding $f : X \to \mathbb{R}^m$ for $m = O(\varepsilon^{-2} \log n)$ such that*

$$\forall x, y \in X,\ (1-\varepsilon)\|x-y\|_2^2 \le \|f(x)-f(y)\|_2^2 \le (1+\varepsilon)\|x-y\|_2^2. \tag{5.1}$$

Lemma 5.0.3 (DJL lemma). *For any $\varepsilon, \delta \in (0, 1/2)$ and integer $d > 1$, there exists a distribution $\mathcal{D}_{\varepsilon,\delta}$ over matrices $\Pi \in \mathbb{R}^{m \times d}$ for $m = O(\varepsilon^{-2} \log(1/\delta))$ such that for any fixed $z \in \mathbb{R}^d$ with $\|z\|_2 = 1$,*

$$\mathop{\mathbb{P}}_{\Pi \sim \mathcal{D}_{\varepsilon,\delta}} \left(\left| \|\Pi z\|_2^2 - 1 \right| > \varepsilon \right) < \delta.$$

مرور

نسبتا خوب می‌شد: فاصله‌ها را تقریبا حفظ کند

$$\forall x, y \in X,\ (1-\varepsilon)\|x-y\|_2^2 \le \|f(x)-f(y)\|_2^2 \le (1+\varepsilon)\|x-y\|_2^2$$

Theorem 5.0.1 (JL lemma [JL84]). *For any $\varepsilon \in (0,1)$ and any $X \subset \mathbb{R}^d$ for $|X| = n$ finite, there exists an embedding $f : X \to \mathbb{R}^m$ for $m = O(\varepsilon^{-2}\log n)$ such that*

$$\forall x, y \in X,\ (1-\varepsilon)\|x-y\|_2^2 \le \|f(x)-f(y)\|_2^2 \le (1+\varepsilon)\|x-y\|_2^2. \tag{5.1}$$

Lemma 5.0.3 (DJL lemma). *For any $\varepsilon, \delta \in (0, 1/2)$ and integer $d > 1$, there exists a distribution $\mathcal{D}_{\varepsilon,\delta}$ over matrices $\Pi \in \mathbb{R}^{m\times d}$ for $m = O(\varepsilon^{-2}\log(1/\delta))$ such that for any fixed $z \in \mathbb{R}^d$ with $\|z\|_2 = 1$,*

$$\mathop{\mathbb{P}}_{\Pi\sim\mathcal{D}_{\varepsilon,\delta}}\left(\left|\|\Pi z\|_2^2 - 1\right| > \varepsilon\right) < \delta.$$

$\Pi_{r,i} = \sigma_{r,i}/\sqrt{m}$ $\sigma_{r,i}$: +۱ یا ــ۱ با احتمال ۱/۲

کران پایین

Johnson–Lindenstrauss Transforms

کران پایین؟

Theorem 5.0.1 (JL lemma [JL84]). *For any $\varepsilon \in (0,1)$ and any $X \subset \mathbb{R}^d$ for $|X| = n$ finite, there exists an embedding $f : X \to \mathbb{R}^m$ for $m = O(\varepsilon^{-2} \log n)$ such that*

$$\forall x, y \in X, \ (1-\varepsilon)\|x-y\|_2^2 \le \|f(x)-f(y)\|_2^2 \le (1+\varepsilon)\|x-y\|_2^2. \tag{5.1}$$

وجود دارد X که هیچ f نیست که ϵ کم باشد

$$X = \{0, e_1, \ldots, e_{n-1}\} \subset \mathbb{R}^n$$

که $\quad f : X \to \mathbb{R}^m$

$$\forall x, y \in X,\ (1-\varepsilon)\|x-y\|_2 \le \|f(x)-f(y)\|_2 \le (1+\varepsilon)\|x-y\|_2$$

$$X = \{0, e_1, \ldots, e_{n-1}\} \subset \mathbb{R}^n$$

که $\quad f : X \to \mathbb{R}^m$

$$\forall x, y \in X,\ (1-\varepsilon)\|x-y\|_2 \leq \|f(x)-f(y)\|_2 \leq (1+\varepsilon)\|x-y\|_2$$

فرض $\quad f(0) = 0 \qquad \tilde{e}_i := f(e_i)$

$$X = \{0, e_1, \dots, e_{n-1}\} \subset \mathbb{R}^n$$

$$f : X \to \mathbb{R}^m$$ که

$$\forall x, y \in X,\ (1-\varepsilon)\|x-y\|_2 \le \|f(x)-f(y)\|_2 \le (1+\varepsilon)\|x-y\|_2$$

فرض $f(0) = 0$ $\quad \tilde{e}_i := f(e_i)$

گام ۱: $\|\tilde{e}_i - 0\|_2 = \|\tilde{e}_i - f(0)\|_2 = (1 \pm \varepsilon)\|e_i\|_2$

$$X = \{0, e_1, \ldots, e_{n-1}\} \subset \mathbb{R}^n$$

که $\quad f : X \to \mathbb{R}^m$

$$\forall x, y \in X,\ (1-\varepsilon)\|x-y\|_2 \le \|f(x) - f(y)\|_2 \le (1+\varepsilon)\|x-y\|_2$$

فرض $\quad f(0) = 0 \qquad \tilde{e}_i := f(e_i)$

گام ۱: $\quad \|\tilde{e}_i - 0\|_2 = \|\tilde{e}_i - f(0)\|_2 = (1 \pm \varepsilon)\|e_i\|_2$

گام ۲: $\quad i \neq j,\ \|\tilde{e}_i - \tilde{e}_j\|_2 \geq (1-\varepsilon)\sqrt{2}$

$$X = \{0, e_1, \ldots, e_{n-1}\} \subset \mathbb{R}^n$$

که $\quad f : X \to \mathbb{R}^m$

$$\forall x, y \in X,\ (1-\varepsilon)\|x-y\|_2 \leq \|f(x) - f(y)\|_2 \leq (1+\varepsilon)\|x-y\|_2$$

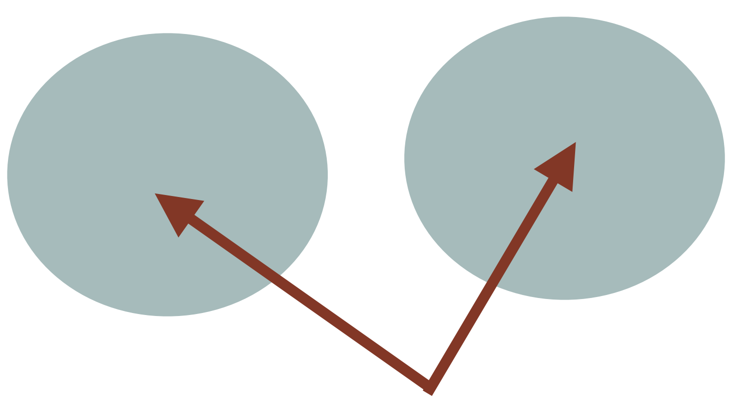
فرض $\quad f(0) = 0 \qquad \tilde{e}_i := f(e_i)$

گام ۱: $\quad \|\tilde{e}_i - 0\|_2 = \|\tilde{e}_i - f(0)\|_2 = (1 \pm \varepsilon)\|e_i\|_2$

گام ۲: $\quad i \neq j,\ \|\tilde{e}_i - \tilde{e}_j\|_2 \geq (1-\varepsilon)\sqrt{2}$

شعاع:

$(1-\varepsilon)\sqrt{2}/2$



$$X = \{0, e_1, \ldots, e_{n-1}\} \subset \mathbb{R}^n$$

که $\quad f : X \to \mathbb{R}^m$

$$\forall x, y \in X,\ (1-\varepsilon)\|x-y\|_2 \leq \|f(x) - f(y)\|_2 \leq (1+\varepsilon)\|x-y\|_2$$

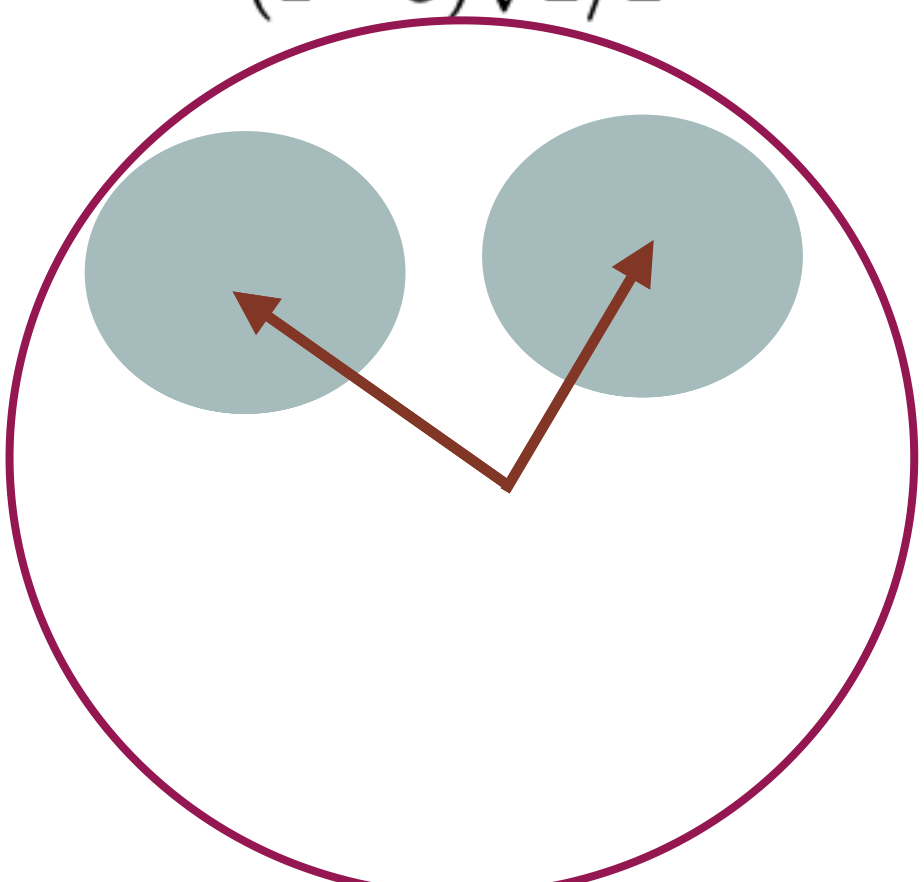
فرض $\quad f(0) = 0 \qquad \tilde{e}_i := f(e_i)$

گام ۱: $\quad \|\tilde{e}_i - 0\|_2 = \|\tilde{e}_i - f(0)\|_2 = (1 \pm \varepsilon)\|e_i\|_2$

گام ۲: $\quad i \neq j,\ \|\tilde{e}_i - \tilde{e}_j\|_2 \geq (1-\varepsilon)\sqrt{2}.$

شعاع:

$(1-\varepsilon)\sqrt{2}/2$

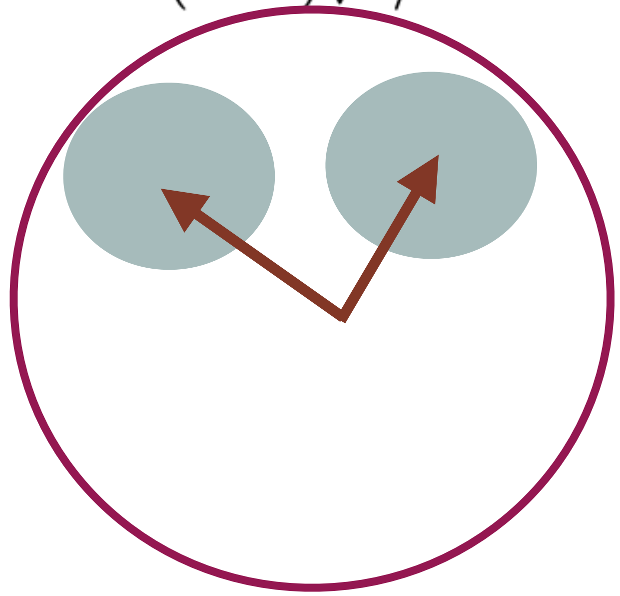


شعاع:

$$(\max_i \|\tilde{e}_i\|_2) + (1-\varepsilon)\sqrt{2}/2$$
$$\leq (1 + \varepsilon + (1-\varepsilon)\sqrt{2}/2)$$

شعاع:

$(1-\varepsilon)\sqrt{2}/2$

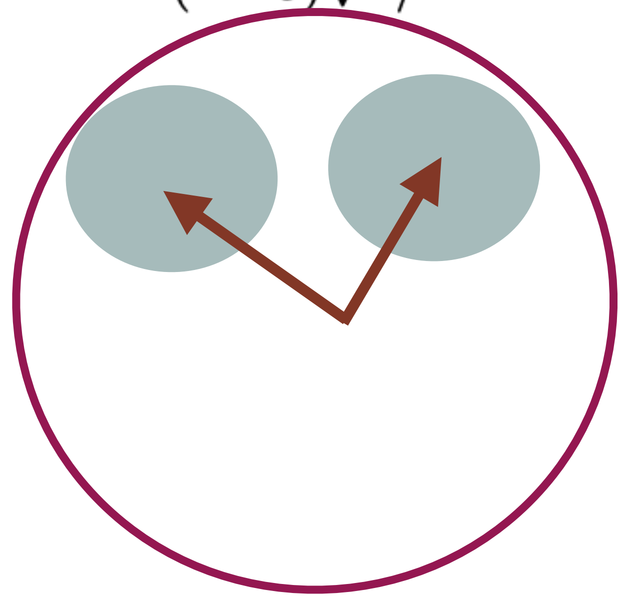


شعاع:

$$(\max_i \|\tilde{e}_i\|_2) + (1-\varepsilon)\sqrt{2}/2$$
$$\leq (1+\varepsilon+(1-\varepsilon)\sqrt{2}/2)$$

شعاع:

$(1-\varepsilon)\sqrt{2}/2$



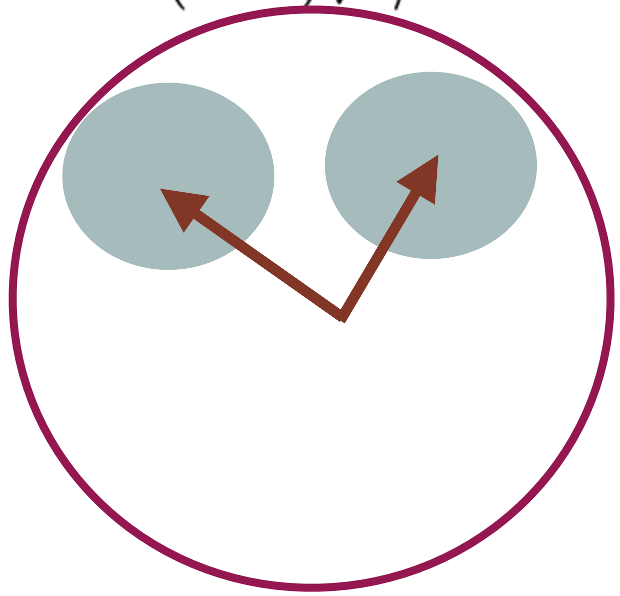
شعاع:

$$(\max_i \|\tilde{e}_i\|_2) + (1-\varepsilon)\sqrt{2}/2$$
$$\leq (1+\varepsilon+(1-\varepsilon)\sqrt{2}/2)$$

$$vol(B_{\ell_2}(0, 1+\varepsilon+(1-\varepsilon)\sqrt{2}/2)) \geq vol(\cup_{i=1}^{n-1} B_{\ell_2}(\tilde{e}_i, (1-\varepsilon)\sqrt{2}/2))$$

شعاع:

$(1-\varepsilon)\sqrt{2}/2$



شعاع:

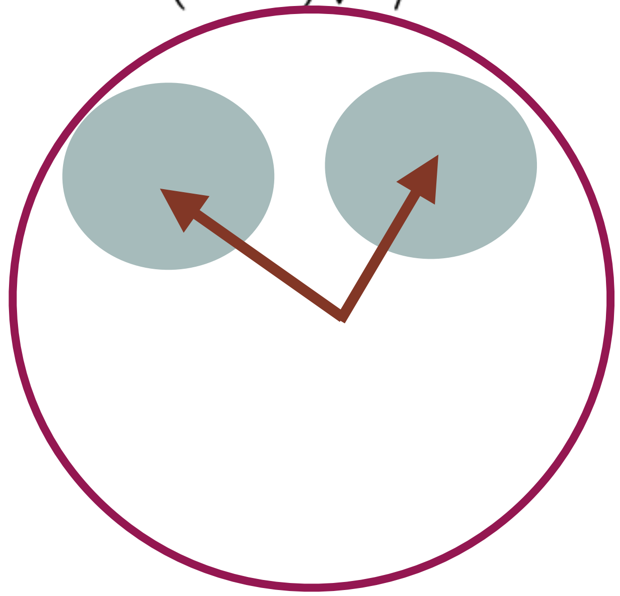
$$(\max_i \|\tilde{e}_i\|_2) + (1-\varepsilon)\sqrt{2}/2 \\ \leq (1+\varepsilon+(1-\varepsilon)\sqrt{2}/2)$$

$$\begin{aligned} vol(B_{\ell_2}(0, 1+\varepsilon+(1-\varepsilon)\sqrt{2}/2)) &\geq vol(\cup_{i=1}^{n-1} B_{\ell_2}(\tilde{e}_i, (1-\varepsilon)\sqrt{2}/2)) \\ &= \sum_{i=1}^{n-1} vol(B_{\ell_2}(\tilde{e}_i, (1-\varepsilon)\sqrt{2}/2)) \end{aligned}$$

عدم اشتراک گوی‌ها

شعاع:

$(1-\varepsilon)\sqrt{2}/2$



شعاع:

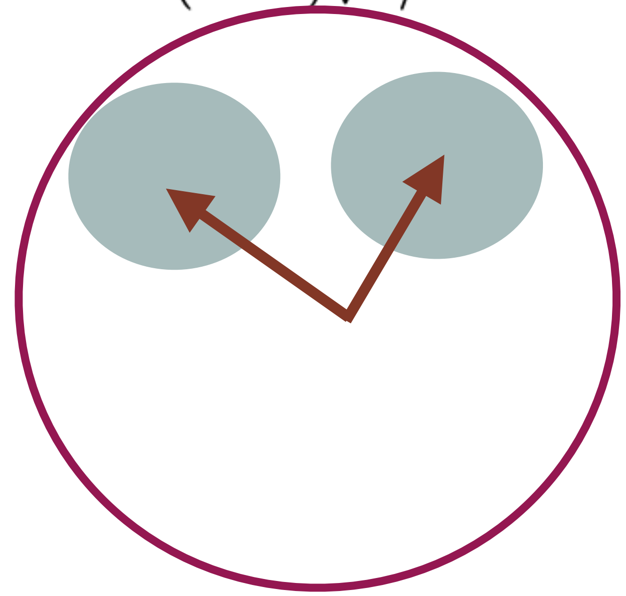
$$(\max_i \|\tilde{e}_i\|_2) + (1-\varepsilon)\sqrt{2}/2 \\ \leq (1+\varepsilon+(1-\varepsilon)\sqrt{2}/2)$$

$$
\begin{aligned}
vol(B_{\ell_2}(0, 1+\varepsilon+(1-\varepsilon)\sqrt{2}/2)) &\geq vol(\cup_{i=1}^{n-1} B_{\ell_2}(\tilde{e}_i, (1-\varepsilon)\sqrt{2}/2)) \\
&= \sum_{i=1}^{n-1} vol(B_{\ell_2}(\tilde{e}_i, (1-\varepsilon)\sqrt{2}/2)) \\
&= (n-1)\cdot vol(B_{\ell_2}(0, (1-\varepsilon)\sqrt{2}/2))
\end{aligned}
$$

عدم اشتراک گوی‌ها

شعاع:

$(1-\varepsilon)\sqrt{2}/2$



شعاع:

$$(\max_i \|\tilde{e}_i\|_2) + (1-\varepsilon)\sqrt{2}/2 \le (1+\varepsilon+(1-\varepsilon)\sqrt{2}/2)$$

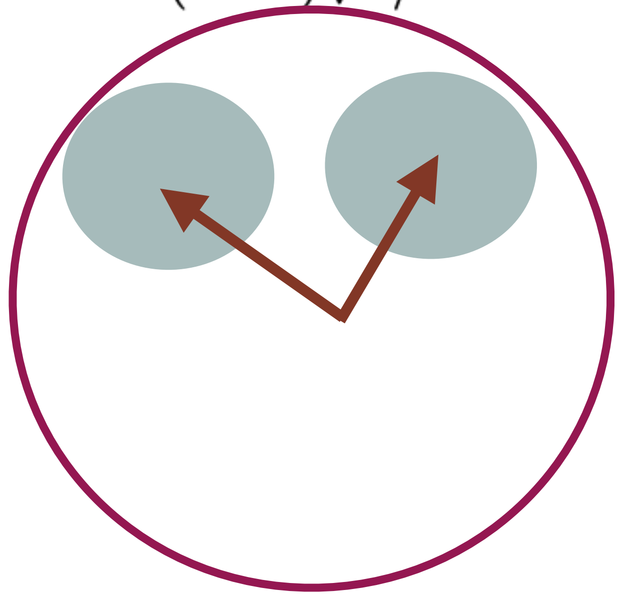
$$\begin{aligned} vol(B_{\ell_2}(0, 1+\varepsilon+(1-\varepsilon)\sqrt{2}/2)) &\geq vol(\cup_{i=1}^{n-1} B_{\ell_2}(\tilde{e}_i, (1-\varepsilon)\sqrt{2}/2)) \\ &= \sum_{i=1}^{n-1} vol(B_{\ell_2}(\tilde{e}_i, (1-\varepsilon)\sqrt{2}/2)) \\ &= (n-1)\cdot vol(B_{\ell_2}(0, (1-\varepsilon)\sqrt{2}/2)) \end{aligned}$$

عدم اشتراک گوی‌ها

$$n-1 \leq \left(\frac{1+\varepsilon+(1-\varepsilon)\sqrt{2}/2}{(1-\varepsilon)\sqrt{2}/2}\right)^m$$

شعاع:

$(1-\varepsilon)\sqrt{2}/2$



شعاع:

$(\max_i \|\tilde{e}_i\|_2) + (1-\varepsilon)\sqrt{2}/2$
$\leq (1+\varepsilon+(1-\varepsilon)\sqrt{2}/2)$

$$vol(B_{\ell_2}(0, 1+\varepsilon+(1-\varepsilon)\sqrt{2}/2)) \geq vol(\cup_{i=1}^{n-1} B_{\ell_2}(\tilde{e}_i, (1-\varepsilon)\sqrt{2}/2))$$

$$= \sum_{i=1}^{n-1} vol(B_{\ell_2}(\tilde{e}_i, (1-\varepsilon)\sqrt{2}/2))$$

$$= (n-1) \cdot vol(B_{\ell_2}(0, (1-\varepsilon)\sqrt{2}/2))$$

عدم اشتراک گوی‌ها

$$n-1 \leq \left(\frac{1+\varepsilon+(1-\varepsilon)\sqrt{2}/2}{(1-\varepsilon)\sqrt{2}/2}\right)^m$$

$$m = \Omega(\log n)$$

برای ϵ ثابت

سریع‌تر

Fast JL Transforms

انگیزش:

- هدف از تبدیل JL:
  - تبدیل بردارها به بعد کم‌تر
  - محاسبات سریع‌تر در بعد کم‌تر

- ماتریس Π:
  - m در d
  - زمان تبدیل: md

- سریع‌تر؟

ایده

- تبدیل Πx —> تبدیل
-

$$\Pi = \frac{1}{\sqrt{m}} SHD$$

ایده

- تبدیل Πx ——> تبدیل
-

هر سطر یک عدد ۱
یکنواخت در ستون‌ها

نمونه‌گیری

$$\Pi = \frac{1}{\sqrt{m}} SHD$$

ایده

- تبدیل Πx ——> تبدیل
-

هر سطر یک عدد ۱
یکنواخت در ستون‌ها

نمونه‌گیری

$$\Pi = \frac{1}{\sqrt{m}} SHD$$

پخش

$$H^\top H = d \cdot I$$

$$\max_{i,j} |H_{i,j}| \leq 1$$

ایده

- تبدیل Πx ——> تبدیل
-

ماتریس قطری با قطر
تصادفی +۱ یا –۱

کانولوشن

نمونه‌گیری

هر سطر یک عدد ۱
یکنواخت در ستون‌ها

$$\Pi = \frac{1}{\sqrt{m}} S H D$$

پخش

$H^\top H = d \cdot I$

$\max_{i,j} |H_{i,j}| \leq 1$

ایده

- تبدیل Πx ——> تبدیل
-

O(m)

ماتریس قطری با قطر تصادفی +۱ یا ــ۱

کانولوشن

نمونه‌گیری

هر سطر یک عدد ۱ یکنواخت در ستون‌ها

$$\Pi = \frac{1}{\sqrt{m}} S H D$$

پخش

$H^\top H = d \cdot I$

$\max_{i,j} |H_{i,j}| \le 1$

ایده

- تبدیل Πx ——> تبدیل
-

$$\Pi = \frac{1}{\sqrt{m}} S H D$$

O(m)

نمونه‌گیری

هر سطر یک عدد ۱
یکنواخت در ستون‌ها

کانولوشن

ماتریس قطری با قطر
تصادفی ۱+ یا ۱ـ

پخش

$H^\top H = d \cdot I$

$\max_{i,j} |H_{i,j}| \le 1$

O(d lg d)

ایده

- تبدیل Πx ——> تبدیل
-

$$\Pi = \frac{1}{\sqrt{m}} S H D$$

هر سطر یک عدد ۱ یکنواخت در ستون‌ها

نمونه‌گیری

O(m)

کانولوشن

O(d)

ماتریس قطری با قطر تصادفی +۱ یا –۱

پخش

O(d lg d)

$H^\top H = d \cdot I$

$\max_{i,j} |H_{i,j}| \leq 1$

Theorem 5.3.2. *Let $z \in \mathbb{R}^d$ be an arbitrary unit norm vector, and suppose $0 < \varepsilon, \delta < 1/2$. Also let $\Pi = \frac{1}{\sqrt{m}} SHD$ as described above with a number of rows equal to $m \gtrsim \varepsilon^{-2} \log(1/\delta) \log(d/\delta)$. Then*

$$\mathbb{P}_{\Pi}(|\|\Pi x\|_2^2 - 1| > \varepsilon) < \delta.$$

Theorem 5.3.2. *Let $z \in \mathbb{R}^d$ be an arbitrary unit norm vector, and suppose $0 < \varepsilon, \delta < 1/2$. Also let $\Pi = \frac{1}{\sqrt{m}} SHD$ as described above with a number of rows equal to $m \gtrsim \varepsilon^{-2} \log(1/\delta) \log(d/\delta)$. Then*

$$\mathbb{P}_{\Pi}(|\|\Pi x\|_2^2 - 1| > \varepsilon) < \delta.$$

مرحله **٤**: $\|HDx\|_\infty \leq \sqrt{2\ln(4d/\delta)}$. با احتمال $1 - \delta/2$

Theorem 5.3.2. *Let $z \in \mathbb{R}^d$ be an arbitrary unit norm vector, and suppose $0 < \varepsilon, \delta < 1/2$. Also let $\Pi = \frac{1}{\sqrt{m}} SHD$ as described above with a number of rows equal to $m \gtrsim \varepsilon^{-2} \log(1/\delta) \log(d/\delta)$. Then*

$$\mathop{\mathbb{P}}_{\Pi}(|\|\Pi x\|_2^2 - 1| > \varepsilon) < \delta.$$

مرحله **ε**: $\|HDx\|_\infty \leq \sqrt{2\ln(4d/\delta)}$ با احتمال $1 - \delta/2$

مرحله ۲: به شرط مرحله **ε**: $(1 - \varepsilon) \leq \|\frac{1}{\sqrt{m}} Sy\|_2^2 \leq (1 + \varepsilon)$ با احتمال $1 - \delta/2$

$$y = HDz$$

مرحله **٤**:

تعریف: $y = HDz$

مرحله **٤**:

تعریف: $y = HDz$

$$y_i = \sum_{j=1}^{d} H_{i,j}\alpha_j z_j = \langle \alpha, H_i z \rangle$$

مرحله ٤:

تعریف: $y = HDz$

$$y_i = \sum_{j=1}^{d} H_{i,j}\alpha_j z_j = \langle \alpha, H_i z \rangle$$

قضیه

$$\forall \lambda > 0,\ \mathbb{P}(|\langle \sigma, x \rangle| > \lambda) \leq 2e^{-\lambda^2/2\|x\|_2^2}$$

$$\mathbb{P}_{\alpha}(|y_i| > \sqrt{2\ln(4d/\delta)}) < 2e^{-\frac{2\ln(4d/\delta)}{2(\sum_{j=1}^{d}|H_{i,j}|^2 z_j^2)}} = \frac{\delta}{2d}$$

مرحله ٤:

تعریف: $y = HDz$

$$y_i = \sum_{j=1}^{d} H_{i,j}\alpha_j z_j = \langle \alpha, H_i z \rangle$$

قضیه

$$\forall \lambda > 0,\ \mathbb{P}(|\langle \sigma, x \rangle| > \lambda) \leq 2e^{-\lambda^2/2\|x\|_2^2}$$

$$\mathbb{P}_{\alpha}(|y_i| > \sqrt{2\ln(4d/\delta)}) < 2e^{-\frac{2\ln(4d/\delta)}{2(\sum_{j=1}^{d} |H_{i,j}|^2 z_j^2)}} = \frac{\delta}{2d}$$

کران اجتماعی:

$$\mathbb{P}_{\alpha}(||y||_\infty > \sqrt{2\ln(4d/\delta)}) < \frac{\delta}{2}$$

$$\|\frac{1}{\sqrt{m}}Sy\|_2^2$$

مرحله ۲:

حکم:

$$\mathop{\mathbb{P}}_{\Pi}(|\|\Pi x\|_2^2 - 1| > \varepsilon) < \delta.$$

$$: j]\frac{1}{m}y^2_{\pi(i)}$$

$$= z D^{T} \overbrace{H^{T} H}^{I} D z$$

$$d \cdot I = d \cdot \|z\|_2^2$$

مرحله ۲ :

حکم:

$$\mathbb{P}_{\Pi}(|\|\Pi x\|_2^2 - 1| > \varepsilon) < \delta.$$

$$\|\frac{1}{\sqrt{m}} Sy\|_2^2 = \sum_{i=1}^{m} \frac{1}{m} y_{\pi(i)}^2$$

$=: X_i$

$: j] \frac{1}{m} y_{\pi(i)}^2$

$= z^T D^T \overbrace{H^T H}^{I} D z$

$\underbrace{H^T H}_{d \cdot I}$

$= d \cdot \|z\|_2^2$

مرحله ۲:

حکم:

$$\mathbb{P}_{\Pi}(|\|\Pi x\|_2^2 - 1| > \varepsilon) < \delta.$$

$$\|\frac{1}{\sqrt{m}}Sy\|_2^2 = \sum_{i=1}^{m} \frac{1}{m} y_{\pi(i)}^2 =: \mathrm{X_i}$$

$$E[X_i] = \sum_{j=1}^{d} P[\pi(i) = j] \frac{1}{m} y_{\pi(i)}^2$$

$$= j]\frac{1}{m}y_{\pi(i)}^2$$

$$= Z D H^T H D Z$$

$$d \cdot I = d \cdot \|Z\|_2^2$$

مرحله ۲:

حکم:

$$\mathbb{P}_{\Pi}(|\|\Pi x\|_2^2 - 1| > \varepsilon) < \delta.$$

$$\left\|\frac{1}{\sqrt{m}}Sy\right\|_2^2 = \sum_{i=1}^{m}\frac{1}{m}y_{\pi(i)}^2 =: X_i$$

$$E[X_i] = \sum_{j=1}^{d}P[\pi(i)=j]\frac{1}{m}y_{\pi(i)}^2$$

$$= \frac{1}{d}\frac{1}{m}||y||_2^2$$

$$j]\frac{1}{m}y_{\pi(i)}^2$$

$$= z^{T}\overbrace{D^{T}\underbrace{H^{T}H}_{d\cdot I}D}^{I}z = d\cdot\|z\|_2^2$$

مرحله ۲:

حکم:

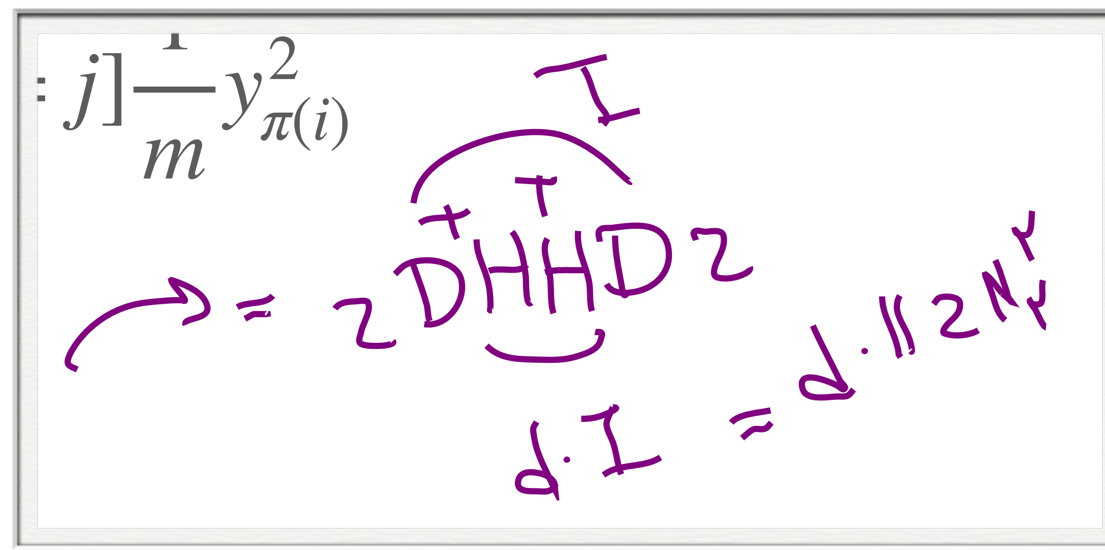
$$\mathbb{P}_{\Pi}(|\|\Pi x\|_2^2 - 1| > \varepsilon) < \delta.$$

$$\|\frac{1}{\sqrt{m}}Sy\|_2^2 = \sum_{i=1}^{m} \frac{1}{m} y_{\pi(i)}^2 =: \mathrm{X_i}$$

$$E[X_i] = \sum_{j=1}^{d} P[\pi(i) = j] \frac{1}{m} y_{\pi(i)}^2$$

$$= \frac{1}{d}\frac{1}{m} ||y||_2^2$$

$$= \frac{1}{m}$$



مرحله ۲:

حکم:

$$\mathop{\mathbb{P}}_{\Pi}(|\|\Pi x\|_2^2 - 1| > \varepsilon) < \delta.$$

$$\|\tfrac{1}{\sqrt{m}} S y\|_2^2 = \sum_{i=1}^{m} \frac{1}{m} y_{\pi(i)}^2 \quad =: X_i$$

$$E[X_i] = \sum_{j=1}^{d} P[\pi(i) = j] \frac{1}{m} y_{\pi(i)}^2$$

$$= \frac{1}{d} \frac{1}{m} ||y||_2^2$$

$$= \frac{1}{m}$$

$$: j] \frac{1}{m} y_{\pi(i)}^2$$

$$= \overbrace{z^{T} D^{T} \underbrace{H H}_{d \cdot I} D}^{I} z = d \cdot \|z\|_2^2$$

$$X := \sum X_i$$

$$E[X] = 1$$

$X := \sum X_i$

$$\mathbb{E}\, X^2 = \sum_i \mathbb{E}\, X_i^2 + \sum_{i \neq j} (\mathbb{E}\, X_i)(\mathbb{E}\, X_j)$$

$$E[X_i^2] = \sum_{j=1}^{d} P[\pi(i) = j] \left(\frac{1}{m} y_{\pi(i)}^2 \right)^2$$

$$\leq \sum_{j=1}^{d} P[\pi(i) = j] \frac{1}{m} y_{\pi(i)}^2 K = \frac{K}{m}$$

$$\leq K + 1$$

مرحله ٤:

$$y_{\pi(i)}^2 \leq 2 \ln(4d/\delta)$$

$$K := \frac{2 \ln(4d/\delta)}{m}$$

$$Var[X] = \mathbb{E}\, X^2 - (\mathbb{E}\, X)^2 \leq K$$

Corollary 1.1.18 (Bernstein's inequality; tail form)**.** *Let $X_1, \ldots, X_n$ be independent, each bounded in magnitude by K almost surely. Write $X := \sum_{i=1}^n X_i$ and define $\sigma^2 := Var[X]$. Then*

$$\forall \lambda > 0, \mathbb{P}(|X - \mathbb{E}\, X| > \lambda) \lesssim e^{-C\lambda^2/\sigma^2} + e^{-C\lambda/K}$$

$$\mathbb{P}\left(\left| \left\| \frac{1}{\sqrt{m}} Sy \right\|_2^2 - 1 \right| > \varepsilon \mid \mathcal{E} \right) \lesssim e^{-C\frac{\varepsilon^2}{K}} + e^{-C\frac{\varepsilon}{K}}$$

Corollary 1.1.18 (Bernstein's inequality; tail form)**.** *Let $X_1, \ldots, X_n$ be independent, each bounded in magnitude by K almost surely. Write $X := \sum_{i=1}^{n} X_i$ and define $\sigma^2 := Var[X]$. Then*

$$\forall \lambda > 0, \mathbb{P}(|X - \mathbb{E}\, X| > \lambda) \lesssim e^{-C\lambda^2/\sigma^2} + e^{-C\lambda/K}$$

$$\begin{aligned}\mathbb{P}\left(\left|\left\|\frac{1}{\sqrt{m}} Sy\right\|_2^2 - 1\right| > \varepsilon \mid \mathcal{E}\right) &\lesssim e^{-C\frac{\varepsilon^2}{K}} + e^{-C\frac{\varepsilon}{K}} \\ &= e^{-C\frac{\varepsilon^2 m}{2\ln(4d/\delta)}} + e^{-C\frac{\varepsilon m}{2\ln(4d/\delta)}}\end{aligned}$$

Corollary 1.1.18 (Bernstein's inequality; tail form). *Let $X_1, \ldots, X_n$ be independent, each bounded in magnitude by K almost surely. Write $X := \sum_{i=1}^{n} X_i$ and define $\sigma^2 := Var[X]$. Then*

$$\forall \lambda > 0, \mathbb{P}(|X - \mathbb{E}\, X| > \lambda) \lesssim e^{-C\lambda^2/\sigma^2} + e^{-C\lambda/K}$$

$$\mathbb{P}\left(\left|\left\|\frac{1}{\sqrt{m}} Sy\right\|_2^2 - 1\right| > \varepsilon \mid \mathcal{E}\right) \lesssim e^{-C\frac{\varepsilon^2}{K}} + e^{-C\frac{\varepsilon}{K}}$$

$$= e^{-C\frac{\varepsilon^2 m}{2\ln(4d/\delta)}} + e^{-C\frac{\varepsilon m}{2\ln(4d/\delta)}}$$

$$\leq \delta/2$$

برای

$$m \gtrsim \varepsilon^{-2} \log(1/\delta) \log(d/\delta)$$

Corollary 1.1.18 (Bernstein's inequality; tail form). *Let $X_1, \dots, X_n$ be independent, each bounded in magnitude by K almost surely. Write $X := \sum_{i=1}^n X_i$ and define $\sigma^2 := Var[X]$. Then*

$$\forall \lambda > 0, \mathbb{P}(|X - \mathbb{E}\, X| > \lambda) \lesssim e^{-C\lambda^2/\sigma^2} + e^{-C\lambda/K}$$

$$\mathbb{P}\left(\left| \left\| \frac{1}{\sqrt{m}} Sy \right\|_2^2 - 1 \right| > \varepsilon \mid \mathcal{E} \right) \lesssim e^{-C\frac{\varepsilon^2}{K}} + e^{-C\frac{\varepsilon}{K}}$$

$$= e^{-C\frac{\varepsilon^2 m}{2\ln(4d/\delta)}} + e^{-C\frac{\varepsilon m}{2\ln(4d/\delta)}}$$

$$\leq \delta/2$$

برای

$$m \gtrsim \varepsilon^{-2} \log(1/\delta) \log(d/\delta)$$

پس:

$$\mathbb{P}\left(\left| \left\| \frac{1}{\sqrt{m}} SHDz \right\|_2^2 - 1 \right| \leq \varepsilon \right) \geq \mathbb{P}(\mathcal{E}) \cdot \mathbb{P}\left(\left| \left\| \frac{1}{\sqrt{m}} SHDz \right\|_2^2 - 1 \right| \leq \varepsilon \mid \mathcal{E} \right) \geq (1 - \delta/2)^2 > 1 - \delta.$$

ضرب کردن ماتریس هادامارد

$$H_1 = [1],$$
$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

$$H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix}$$