Scalable ML

10605-10805

Johnson-Lindenstrauss Lemma

Barnabás Póczos

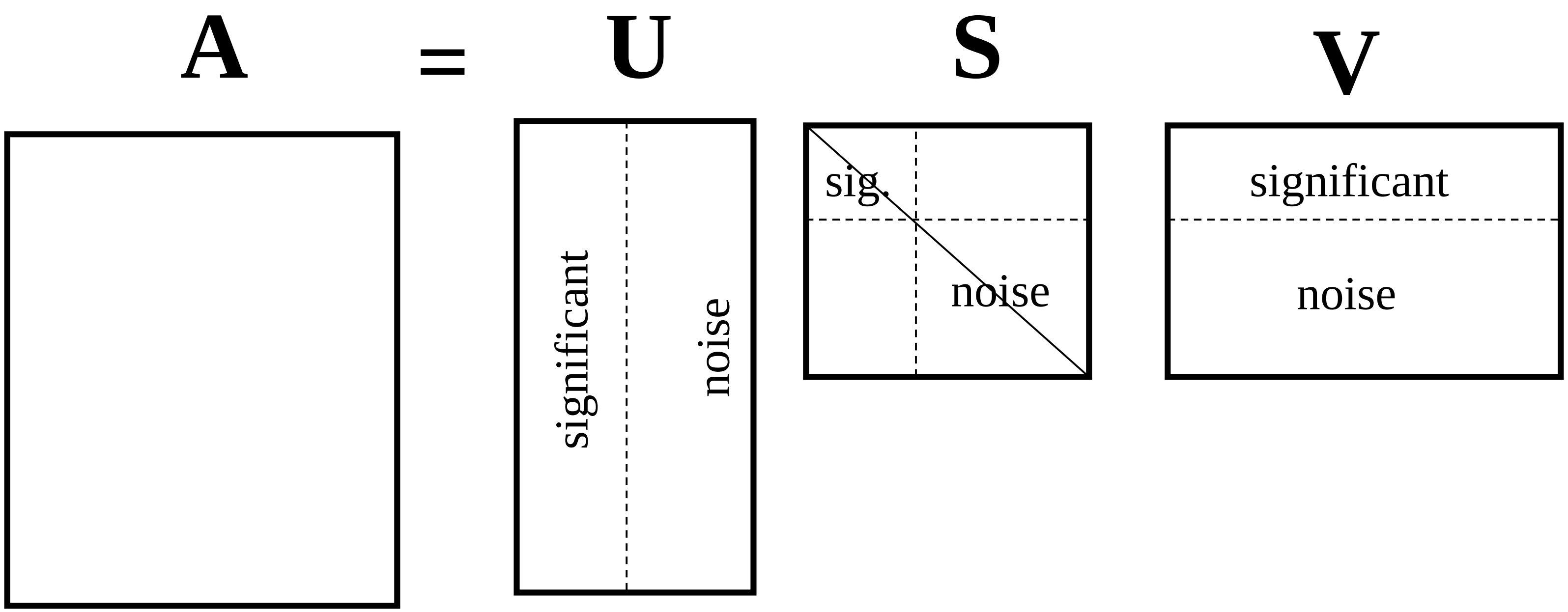
Dimension reduction with SVD

Singular Value Decomposition of the data matrix **A**.

$$\text{Let } A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1d} \\ a_{21} & a_{22} & \cdots & a_{2d} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nd} \end{pmatrix} \in \mathbb{R}^{n \times d}$$

n: num of instances,
d: dimension

A = USV

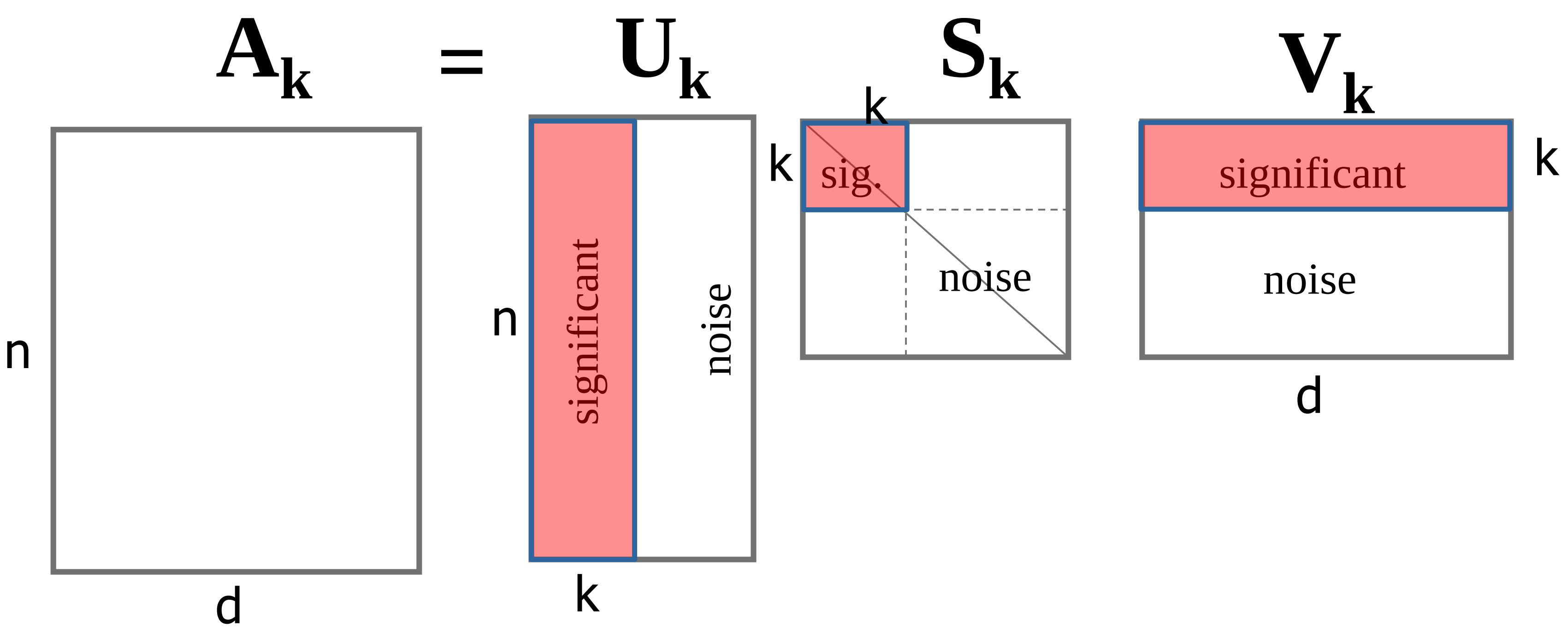


Dim reduction with SVD

$$\text{Let } A_k = \begin{pmatrix} \hat{a}_{11} & \hat{a}_{12} & \cdots & \hat{a}_{1d} \\ \hat{a}_{21} & \hat{a}_{22} & \cdots & \hat{a}_{2d} \\ \vdots & \vdots & \cdots & \vdots \\ \hat{a}_{n1} & \hat{a}_{n2} & \cdots & \hat{a}_{nd} \end{pmatrix} \in \mathbb{R}^{n \times d}$$

be a rank k approximation given by SVD

A = USV $\quad$ **$A_k = U_k S_k V_k$**



Dim reduction with SVD

Lemma [SVD provides the best rank k approximation]

$$\|A - A_k\|_{Fro} \leq \|A - D\|_{Fro} \quad \forall D \in \mathbb{R}^{n \times d} \text{ rank } k \text{ matrices}$$

$$\|A - A_k\|_2 \leq \|A - D\|_2 \quad \forall D \in \mathbb{R}^{n \times d} \text{ rank } k \text{ matrices}$$

Proof [out of scope]

Issues with SVD

Observation:

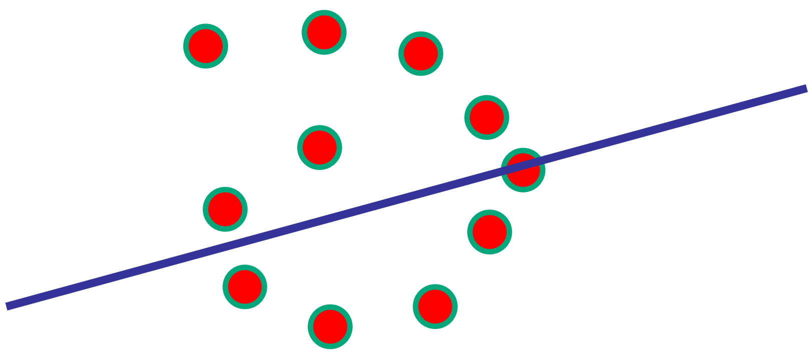
Although SVD provides the best mtx approximation globally, it might ruin local structures:

Data points which were far might get very close after projection with SVD

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \vdots & \vdots \\ a_{n1} & a_{n2} \end{pmatrix} \in \mathbb{R}^{n \times 2}$$

Let $k = 1$

$$A_k = \begin{pmatrix} \hat{a}_{11} & \hat{a}_{12} \\ \hat{a}_{21} & \hat{a}_{22} \\ \vdots & \vdots \\ \hat{a}_{n1} & \hat{a}_{n2} \end{pmatrix} = \begin{pmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{n1} \end{pmatrix} S_{11} \begin{pmatrix} v_{11} & v_{12} \end{pmatrix}$$

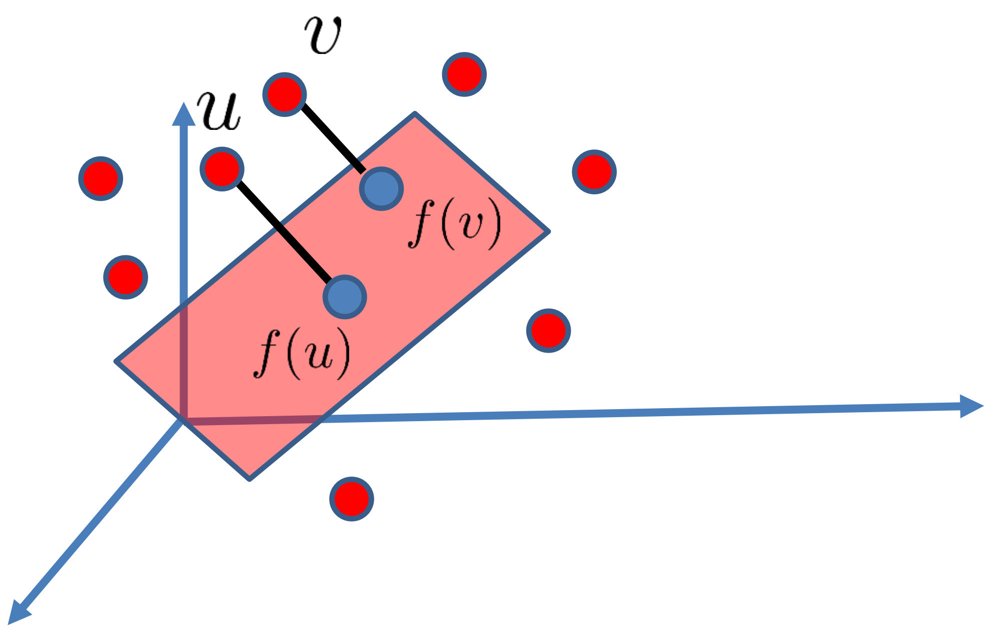


How can we fix it?

Johnson-Lindenstrauss Lemma

Informal theorem

Any n points in $\mathbb{R}^d$ can be linearly projected into a $O(\log(n)/\epsilon^2)$ dim subspace without distorting the pairwise distances more than a $(1 \pm \epsilon)$ factor.



$$(1-\epsilon)\|u-v\|^2 \leq \|f(u)-f(v)\|^2 \leq (1+\epsilon)\|u-v\|^2, \quad \forall u, v \in V$$

Johnson-Lindenstrauss Lemma

Theorem 1 [The Johnson-Lindenstrauss Lemma]

Let $0 < \epsilon < 1$

Let $k \geq \frac{4\log(n)}{\epsilon^2 - \epsilon^3/3}$

Let $V = \{v_1, \ldots, v_n\}$ be an arbitrary set of n points in $\mathbb{R}^d$. $(v_i \in \mathbb{R}^d)$

Then there exists a map $f\colon \mathbb{R}^d \to S$, such that

f is a linear projection,

S is a k-dim subspace through the origin in R^d,

$$(1-\epsilon)\|u-v\|^2 \leq \|f(u)-f(v)\|^2 \leq (1+\epsilon)\|u-v\|^2, \quad \forall u, v \in V$$

Furthermore, this function f can be found with a randomized algorithm

Significance

- Using the JL Lemma, we can represent a dataset with a smaller dimensional dataset while keeping the pairwise distances mostly unchanged

- JL Lemma can speed up algorithms whose running time suffer from high-dimensions

Tightness

Theorem 2a [The JL Lemma is essentially tight]

If we are given a set of n points in $\mathbb{R}^d$: $V = \{v_1, \ldots, v_n\}$, $v_i \in \mathbb{R}^d$

such that the pairwise disntances $(\{\|v_i - v_j\|^2\})$ are all between $[1-\epsilon, 1+\epsilon] \quad \forall 1 \leq i \neq j \leq n$

then V requires at least

$$\Omega\left(\frac{\log n}{\epsilon^2 \log(1/\epsilon)}\right) \text{ dimension}$$

Tightness

Theorem 2b [The JL Lemma is tight]

Larsen, Nelson 2017

For any $d, n \geq 2$ and $\frac{1}{(\min\{n,d\})^{0.49999}} < \epsilon < 1$

there exists a set of n vectors $V = \{v_1, \ldots, v_n\} \subset\in \mathbb{R}^d$

such that for any embedding $f : V \to \mathbb{R}^m$

satisfying

$$(1-\epsilon)\|u-v\|^2 \leq \|f(u)-f(v)\|^2 \leq (1+\epsilon)\|u-v\|^2, \quad \forall u, v \in V$$

must have at least

$$m = \Omega\left(\frac{\log n}{\epsilon^2}\right) \text{dimension}$$

Proofs

Proof Idea

Let $1 \leq i, j \leq n$ be fixed.

We will prove that if f is a random projection to a k-dim subspace, then

$$Pr\left(\frac{\|f(v_i) - f(v_j)\|^2}{\|v_i - v_j\|^2} \notin [1-\epsilon, 1+\epsilon]\right) \leq \frac{2}{n^2}$$

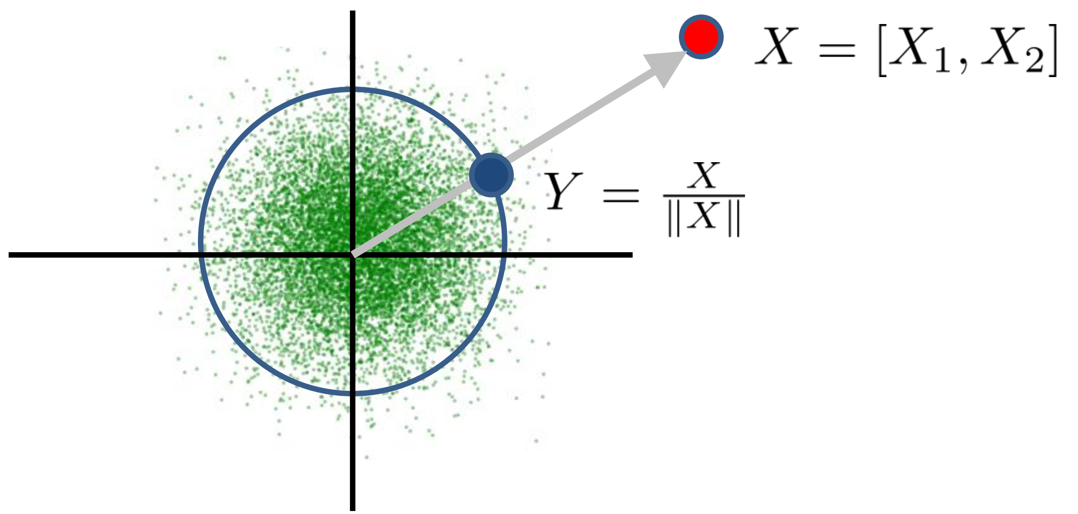
- Then use union bound to have a bound for all i,j
- Try several random projections till we are happy with the embeddings

Preliminaries

Let $X = [X_1, X_2, \ldots, X_d]$ be d indpendent $\mathcal{N}(0,1)$ random variables.

Let $Y = \frac{X}{\|X\|} \in \mathbb{R}^d$

Observation: Y is uniformly distributed on the surface of a d-dim unit sphere



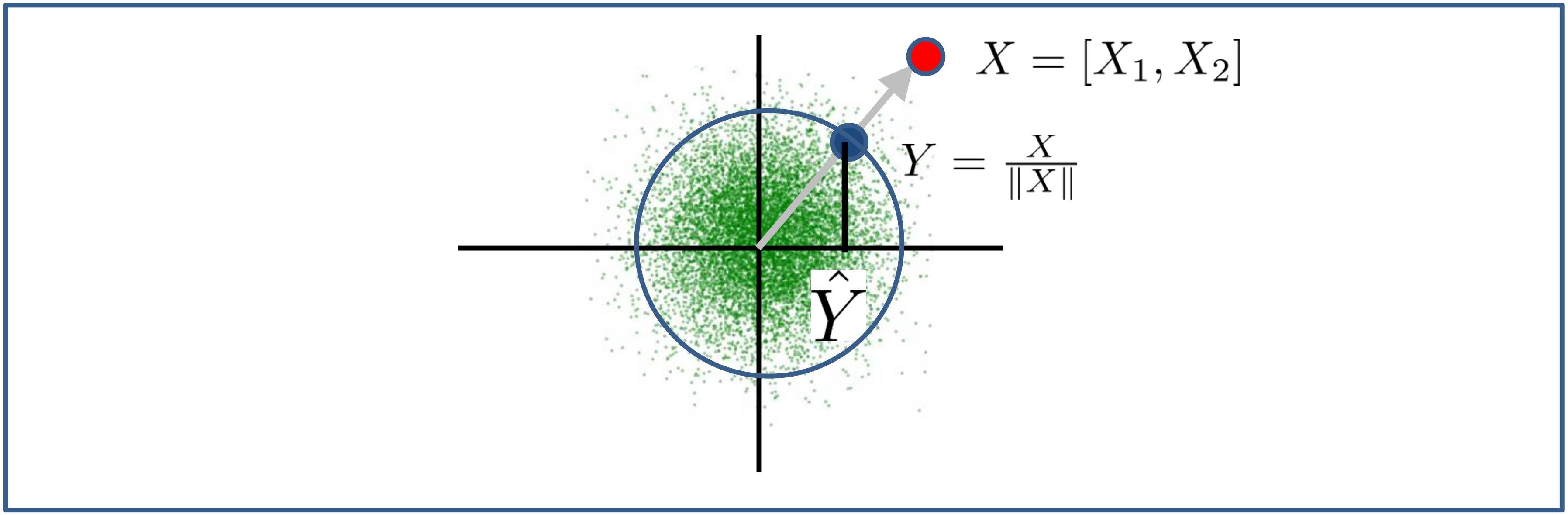
Preliminaries

Lemma 1 [Projection to 1st k coordinates, expected squared length]

Let $\hat{Y}$ be the projection of Y onto its first k cooridnates

Let $L = L(Y) = \|\hat{Y}\|^2$

Observation: $\mathbb{E}[L] = \mathbb{E}[\|\hat{Y}\|^2] = \frac{k}{d}$



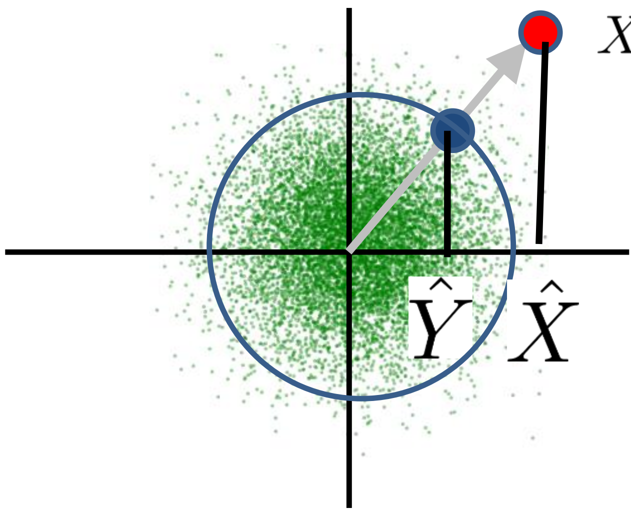
Length of a random projection

Corollary 1 [Generalization to arbitrary vectors]

Let $\hat{X}$ be the projection of X onto its first k cooridnates

Let $L = L(X) = \|\hat{X}\|^2$

Observation: $\mathbb{E}[L] = \mathbb{E}[\|\hat{X}\|^2] = \frac{k}{d}\|X\|^2$



$$X = [X_1, X_2]$$

$$Y = \frac{X}{\|X\|}$$

$$Y = \frac{X}{\|X\|}$$

$$\hat{Y} = \frac{\hat{X}}{\|X\|}$$

Length of a random projection

Lemma 2
[Projection to 1st k coordinates, deviation from expected squared length]

Let $k < d$.

If $\beta < 1 \Rightarrow$

$$Pr\left[\|\hat{Y}\|^2 \leq \beta\frac{k}{d}\right] \leq \beta^{k/2}\left(1 + \frac{(1-\beta)k}{d-k}\right)^{(d-k)/2}$$
$$\leq \exp\left(\frac{k}{2}(1-\beta+\log\beta)\right)$$

If $\beta > 1 \Rightarrow$

$$Pr\left[\|\hat{Y}\|^2 \geq \beta\frac{k}{d}\right] \leq \beta^{k/2}\left(1 + \frac{(1-\beta)k}{d-k}\right)^{(d-k)/2}$$
$$\leq \exp\left(\frac{k}{2}(1-\beta+\log\beta)\right)$$

Informal meaning: $Pr\left[\|\hat{Y}\|^2 \text{ is far from its mean}\right] \leq \text{ small}$

Length of a random projection

Corollary 2a [Generalization to arbitrary vectors]

Let $\hat{X}$ be the projection of X onto its fist k cooridnates

Let $L = \|\hat{X}\|^2$

If $\beta < 1 \Rightarrow$

$$Pr\left[\|\hat{Y}\|^2 \leq \beta\frac{k}{d}\right] = Pr\left[\|\hat{Y}\|^2\|X\|^2 \leq \beta\frac{k}{d}\|X\|^2\right]$$

$$= Pr\left[\|\hat{X}\|^2 \leq \beta\frac{k}{d}\|X\|^2\right]$$

$$\hat{Y} = \frac{\hat{X}}{\|X\|} \Rightarrow \hat{Y}\|X\| = \hat{X}$$

Therefore,

$$\Rightarrow Pr\left[\|\hat{X}\|^2 \leq \beta\frac{k}{d}\|X\|^2\right] \leq \beta^{k/2}\left(1 + \frac{(1-\beta)k}{d-k}\right)^{(d-k)/2}$$

$$\leq \exp\left(\frac{k}{2}(1-\beta+\log\beta)\right)$$

Length of a random projection

Corollary 2b [Generalization to arbitrary vectors]

Similarly,

If $\beta > 1 \Rightarrow$

$$Pr\left[\|\hat{Y}\|^2 \geq \beta\frac{k}{d}\right] = Pr\left[\|\hat{Y}\|^2\|X\|^2 \geq \beta\frac{k}{d}\|X\|^2\right]$$
$$= Pr\left[\|\hat{X}\|^2 \geq \beta\frac{k}{d}\|X\|^2\right]$$

Therefore,

$$\Rightarrow Pr\left[\|\hat{X}\|^2 \geq \beta\frac{k}{d}\|X\|^2\right] \leq \beta^{k/2}\left(1 + \frac{(1-\beta)k}{d-k}\right)^{(d-k)/2}$$
$$\leq \exp\left(\frac{k}{2}(1 - \beta + \log\beta)\right)$$

Random vector vs Random subspace

Projection of a random vector to its first k coordinates

vs

Projection of a fixed vector to a random subspace

It doesn't matter if

- we project a "uniform direction" random vector to its first k coordinates

or

- we project a given vector to a "uniform direction" k-dim random subspace

$$\text{The distribution of } \| \text{ projected vector } \|^2 \text{ is the same.}$$

Proof of the JL Lemma

If $d \leq k \Rightarrow$ Theorem 1 is trivial.

If $d \geq k \Rightarrow$ fix i and j.

Let S be a random subspace
(uniform direction, origin is in the subspace)

Let $\hat{v}_i, \hat{v}_j$ be the projection of v_i, $v_j \in V$ into S.

Let $L = \|\hat{v}_i - \hat{v}_j\|^2$

Let $\mu = \frac{k}{d}\|v_i - v_j\|^2$

Observation: Using Corollary 1, we have that

$$\begin{aligned}\mathbf{E}[L] &= \mathbf{E}[\|\hat{v}_i - \hat{v}_j\|^2] \\ &= \frac{k}{d}\|v_i - v_j\|^2\end{aligned}$$

Proof of the JL Lemma (continued)

Let $\beta = 1 - \epsilon$.

Using the above notation, we can apply Corollary 2a:

$$Pr\left[\|\hat{X}\|^2 \leq \beta \frac{k}{d} \|X\|^2\right] \leq \exp\left(\frac{k}{2}(1 - \beta + \log \beta)\right)$$

$$Pr\left[\overbrace{L}^{\|\hat{X}\|^2} \leq \overbrace{(1-\epsilon)}^{\beta} \frac{k}{d} \overbrace{\|v_i - v_j\|^2}^{\|X\|^2}\right] \leq \exp\left(\frac{k}{2}\Big(\underbrace{1-(1-\epsilon)}_{\epsilon} + \underbrace{\log(1-\epsilon)}_{\leq -\epsilon - \epsilon^2/2}\Big)\right)$$

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

$$\leq \exp\left(\frac{-k}{4}\epsilon^2\right)$$

$$\frac{-k}{4}\epsilon^2 \leq (-2\log n)$$

$$k \geq \frac{8 \log n}{\epsilon^2}$$

This hold under the condition of JS Lemma

$$\leq \exp(-2\log n) = \frac{1}{n^2} \qquad (*1)$$

Proof of the JL Lemma (continued)

Similarly, Let $\beta = 1 + \epsilon$.

Using the above notation, we can apply Corollary 2b:

$$Pr\left[\|\hat{X}\|^2 \geq \beta \frac{k}{d}\|X\|^2\right] \leq \exp\left(\frac{k}{2}(1 - \beta + \log \beta)\right)$$

$$Pr\left[\overbrace{L}^{\|\hat{X}\|^2} \geq \overbrace{(1+\epsilon)}^{\beta} \frac{k}{d} \overbrace{\|v_i - v_j\|^2}^{\|X\|^2}\right] \leq \exp\left(\frac{k}{2}\left(\underbrace{1 - (1+\epsilon)}_{-\epsilon} + \underbrace{\log(1+\epsilon)}_{\leq \epsilon - \epsilon^2/2 + \epsilon^3/3}\right)\right)$$

$$\leq \exp\left(-\frac{k}{2}\left(\epsilon^2/2 - \epsilon^3/3\right)\right)$$

$$-\frac{k}{2}\left(\epsilon^2/2 - \epsilon^3/3\right) \leq -2\log n$$

$$4\log n \leq k(\epsilon^2/2 - \epsilon^3/3)$$

$$\frac{4\log n}{\epsilon^2/2 - \epsilon^3/3} \leq k$$

$$\leq \exp\left(-2\log n\right) = \frac{1}{n^2} \qquad (*2)$$

This hold under the condition of JS Lemma

The Projection Map

Let the map $f(v_i) \doteq \sqrt{\frac{d}{k}}\hat{v}_i$, for all $1 \le i \le n$

$$\Rightarrow \|f(v_i) - f(v_j)\|^2 = \|\sqrt{\frac{d}{k}}\hat{v}_i - \sqrt{\frac{d}{k}}\hat{v}_j\|^2$$
$$= \frac{d}{k}\|\hat{v}_i - \hat{v}_j\|^2$$

Therefore,

$$Pr\left(\frac{\|f(v_i) - f(v_j)\|^2}{\|v_i - v_j\|^2} \ge 1 + \epsilon\right) = Pr\left(\frac{d}{k}\frac{\|\hat{v}_i - \hat{v}_j)\|^2}{\|v_i - v_j\|^2} \ge 1 + \epsilon\right)$$
$$= Pr\left(\|\hat{v}_i - \hat{v}_j)\|^2 \ge (1+\epsilon)\frac{k}{d}\|v_i - v_j\|^2\right)$$

(*3) $\le \dfrac{1}{n^2}$ From (*2)

The Projection Map

Similarly,

$$Pr\left(\frac{\|f(v_i) - f(v_j)\|^2}{\|v_i - v_j\|^2} \leq 1 - \epsilon\right) = Pr\left(\frac{d}{k}\frac{\|\hat{v}_i - \hat{v}_j)\|^2}{\|v_i - v_j\|^2} \leq 1 - \epsilon\right)$$

$$= Pr\left(\|\hat{v}_i - \hat{v}_j)\|^2 \leq (1-\epsilon)\frac{d}{k}\|v_i - v_j\|^2\right)$$

(*4) $$\leq \frac{1}{n^2}$$ From (*1)

Proof of the JL Lemma (continued)

From (*3) and (*4), using the union bound:

For a fixed i, and j $(1 \leq i, j \leq n)$, we have that

$$Pr\left(\frac{\|f(v_i) - f(v_j)\|^2}{\|v_i - v_j\|^2} \notin [1-\epsilon, 1+\epsilon]\right) \leq \frac{1}{n^2} + \frac{1}{n^2} = \frac{2}{n^2}$$

Therefore,

$$\begin{aligned} Pr\left(\exists i, j : \frac{\|f(v_i) - f(v_j)\|^2}{\|v_i - v_j\|^2} \notin [1-\epsilon, 1+\epsilon]\right) &\leq \binom{n}{2}\frac{2}{n^2} \\ &= \frac{n(n-1)}{2}\frac{2}{n^2} \\ &= \frac{n-1}{n} \\ &= 1 - \frac{1}{n} \end{aligned}$$

Proof of the JL Lemma (continued)

Therefore,

$$Pr\left(\forall i,j : \frac{\|f(v_i) - f(v_j)\|^2}{\|v_i - v_j\|^2} \in [1-\epsilon, 1+\epsilon]\right) \geq= \frac{1}{n}$$

If we want the r.h.s to be $> 1-\delta$, we just need to generate multiple $(= m)$ independent embeddings.

One of them will be good by high probability.

We need $(1-\frac{1}{n})^m < \delta$ $\quad\Leftrightarrow m\log(1-\frac{1}{n}) < \log\delta$

(failure probability)

$$\Leftrightarrow m > \frac{\log\delta}{\log(1-\frac{1}{n})}$$

Since the failure probability can be arbitrarily close to 0, this proves the JS Lemma.

Proof of the JL Lemma (continued)

All that left is to prove Lemma 2, that is

$$\text{If } \beta < 1 \Rightarrow \quad Pr\left[\|\hat{X}\|^2 \leq \beta\frac{k}{d}\|X\|^2\right] \leq \beta^{k/2}\left(1+\frac{(1-\beta)k}{d-k}\right)^{(d-k)/2}$$

(*5)

$$\leq \exp\left(\frac{k}{2}(1-\beta+\log\beta)\right)$$

$$\text{If } \beta > 1 \Rightarrow \quad Pr\left[\|\hat{X}\|^2 \geq \beta\frac{k}{d}\|X\|^2\right] \leq \beta^{k/2}\left(1+\frac{(1-\beta)k}{d-k}\right)^{(d-k)/2}$$

(*6)

$$\leq \exp\left(\frac{k}{2}(1-\beta+\log\beta)\right)$$

Proof of the JL Lemma (continued)

Proof:

$$
\begin{aligned}
Pr\left[\|\hat{X}\|^2 \le \beta\frac{k}{d}\|X\|^2\right] &= Pr\left[d\|\hat{X}\|^2 \le \beta k\|X\|^2\right] \\
&= Pr\left[d(X_1^2 + \ldots + X_k^2) \le \beta k(X_1^2 + \ldots + X_d^2)\right] \\
&= Pr\left[\beta k(X_1^2 + \ldots + X_d^2) - d(X_1^2 + \ldots + X_k^2) \ge 0\right] \\
&= Pr\left[\exp\left\{t\Big(\beta k(X_1^2 + \ldots + X_d^2) - d(X_1^2 + \ldots + X_k^2)\Big)\right\} \ge 1\right] \quad \forall t > 0
\end{aligned}
$$

Lemma: [Markov's inequality]

$$
Pr(Z \ge a) \le \frac{\mathbb{E}(Z)}{a}, \quad \forall Z, a \ge 0
$$

Therefore,

$$
\begin{aligned}
Pr\left[\|\hat{X}\|^2 \le \beta\frac{k}{d}\|X\|^2\right] &= Pr\left[\overbrace{\exp\left\{t\Big(\beta k(X_1^2 + \ldots + X_d^2) - d(X_1^2 + \ldots + X_k^2)\Big)\right\}}^{Z} \ge 1\right] \quad \forall t > 0 \\
&\le \mathbb{E}\left[\exp\left\{t\Big(\beta k(X_1^2 + \ldots + X_d^2) - d(X_1^2 + \ldots + X_k^2)\Big)\right\}\right]
\end{aligned}
$$

Proof of the JL Lemma (continued)

This is what we know so far:

$$Pr\left[\|\hat{X}\|^2 \leq \beta\frac{k}{d}\|X\|^2\right] \leq \mathbb{E}\left[\exp\left\{t\Big(\beta k(X_1^2+\ldots+X_d^2) - d(X_1^2+\ldots+X_k^2)\Big)\right\}\right]$$

$$= \mathbb{E}\left[\exp\left\{(t\beta k - td)\sum_{j=1}^{k} X_j^2 + \beta kt \sum_{j=k+1}^{d} X_j^2\right\}\right]$$

$$= \mathbb{E}\left[\prod_{j=1}^{k}\exp\left\{(t\beta k - td)X_j^2\right\} \prod_{j=k+1}^{d}\exp\left\{\beta kt X_j^2\right\}\right]$$

$$= \prod_{j=1}^{k}\mathbb{E}\left[\exp\left\{(t\beta k - td)X_j^2\right\}\right] \prod_{j=k+1}^{d}\mathbb{E}\left[\exp\left\{\beta kt X_j^2\right\}\right]$$

$$= \left(\mathbb{E}\left[\exp\left\{(t\beta k - td)Z^2\right\}\right]\right)^k \left(\mathbb{E}\left[\exp\left\{\beta kt Z^2\right\}\right]\right)^{d-k}$$

Where $Z \sim \mathcal{N}(0,1)$

Proof of the JL Lemma (continued)

Lemma [Moment generating function of X²]

$$\text{Let } X \sim \mathcal{N}(0,1). \Rightarrow \mathbb{E}[\exp(sX^2)] = \frac{1}{\sqrt{1-2s}} = g(s), \forall -\infty < s < 1/2$$

Proof: Out of scope

$$\text{Where } Z \sim \mathcal{N}(0,1)$$

$$Pr\left[\|\hat{X}\|^2 \leq \beta \frac{k}{d}\|X\|^2\right] \leq \left(\mathbb{E}\left[\exp\left\{(t\beta k - td)Z^2\right\}\right]\right)^k \left(\mathbb{E}\left[\exp\left\{\beta k t Z^2\right\}\right]\right)^{d-k}$$

$$= g(t\beta k - td)^k g(t\beta k)^{d-k}$$

$$= (1 - 2t\beta k - 2td)^{-k/2}(1 - 2t\beta k)^{(k-d)/2}$$

$$\text{For all } t \text{ such that } (1 - 2t\beta k - 2td) > 0 \text{ and } (1 - 2t\beta k) > 0.$$

If we minimize the r.h.s. in *t* we get (*5), what we wanted to prove. Q.E.D

Proof of the JL Lemma (continued)

All that left is to prove (*6):

$$\text{If } \beta > 1 \Rightarrow \quad Pr\left[\|\hat{X}\|^2 \geq \beta\frac{k}{d}\|X\|^2\right] \leq \beta^{k/2}\left(1 + \frac{(1-\beta)k}{d-k}\right)^{(d-k)/2}$$

$$\leq \exp\left(\frac{k}{2}(1-\beta+\log\beta)\right)$$

Its proof is similar to (*5). QED

Issues with the JL Lemma

- The JL Algorithm projects the points of the dataset onto a random hyperplane through the origin.

- This projection might be computationally expensive.

- **Goal**: Design a new algorithm where the random projections in JL are replaced with much simpler operations.

Database Friendly Random Projections

Let $V = \{v_1, \ldots, v_n\}$ be an arbitrary set of n points in $\mathbb{R}^d$. $(v_i \in \mathbb{R}^d)$

Let $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1d} \\ a_{21} & a_{22} & \cdots & a_{2d} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nd} \end{pmatrix} \in \mathbb{R}^{n \times d}$

the representation of the n points

The i^{th} row is v_i

Let $0 < \epsilon < 1$

Let $\beta > 0$

Let $k_0 = \frac{4+2\beta}{\epsilon^2 - \epsilon^3/3} \log(n)$

Let $k > k_0$

Database Friendly Random Projections

Let $R = \begin{pmatrix} r_{11} & r_{r2} & \dots & a_{1k} \\ r_{21} & r_{22} & \dots & r_{2d} \\ \vdots & \vdots & \dots & \vdots \\ r_{d1} & r_{d2} & \dots & r_{dk} \end{pmatrix} \in \mathbb{R}^{d \times k}$ be a random matrix such that

$R(i,j) = r_{ij}$, where

$$r_{i,j} = \begin{cases} 1 & \text{with prob } 1/2 \\ -1 & \text{with prob } 1/2 \end{cases}$$

or

$$r_{i,j} = \sqrt{3} \begin{cases} 1 & \text{with prob } 1/6 \\ 0 & \text{with prob } 2/3 \\ -1 & \text{with prob } 1/6 \end{cases}$$

Database Friendly Random Projections

Let $E = \frac{1}{\sqrt{k}} AR$

$$= \frac{1}{\sqrt{k}} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ a_{21} & a_{22} & \dots & a_{2d} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nd} \end{pmatrix} \begin{pmatrix} r_{11} & r_{r2} & \dots & a_{1k} \\ r_{21} & r_{22} & \dots & r_{2d} \\ \vdots & \vdots & \dots & \vdots \\ r_{d1} & r_{d2} & \dots & r_{dk} \end{pmatrix} \in \mathbb{R}^{n \times k}$$

$$= \begin{pmatrix} e_{11} & e_{12} & \dots & e_{1k} \\ e_{21} & e_{22} & \dots & e_{2k} \\ \vdots & \vdots & \dots & \vdots \\ e_{n1} & e_{n2} & \dots & e_{nk} \end{pmatrix}$$

Let $f : V \to \mathbb{R}^k$ map the i^{th} row of A to the i^{th} row of E.

Database Friendly Random Projections

Under these conditions,

with probability at least $(1 - n^{-\beta})$

$$(1-\epsilon)\|u-v\|^2 \leq \|f(u)-f(v)\| \leq (1+\epsilon)\|u-v\|^2, \quad \forall u, v \in V$$

All operations to cacluate $E = \frac{1}{\sqrt{k}} AR$ are easy to implement

since $r_{ij} \in \{-1, +1\}$ or $r_{ij} \in \{-\sqrt{3}, 0 + \sqrt{3}\}$

Thanks for your Attention! ☺