We recall, if we have γ_n, standard Gaussian measure in n dimensions, and take $\Phi : \mathbb{R}^n \to S^{n-1}$, $x \mapsto \frac{x}{\|x\|}$; then the image measure induced by Φ is the normalized surface measure.

<u>Cor</u> If μ_n is rot. inv. prob. measure on S^{n-1}, $V \subset \mathbb{R}^n$ a k-dim subspace of $\mathbb{R}^n$ and P assoc. orth. proj. onto V, then

$$\mu_n\left(\left\{x \in S^{n-1} : \sqrt{\tfrac{n}{k}}\,\|Px\| \geq \frac{1}{1-\varepsilon}\right\}\right) \leq e^{-\varepsilon^2 n/4} + e^{-\varepsilon^2 k/4}$$

and

$$\mu_n\left(\left\{x \in S^{n-1} : \sqrt{\tfrac{n}{k}}\,\|Px\| \leq 1-\varepsilon\right\}\right) \leq e^{-\varepsilon^2 n/4} + e^{-\varepsilon^2 k/4}$$

Pf. Consider

$$\mu_n(\{ y \in S^{n-1} : \sqrt{\tfrac{n}{k}} \|Py\| \geq \tfrac{1}{1-\varepsilon} \})$$

$$= \gamma_n(\{ x \in \mathbb{R}^n : \sqrt{\tfrac{n}{k}} \| P \underbrace{\tfrac{x}{\|x\|}}_{\phi(x)} \| \geq \tfrac{1}{1-\varepsilon} \})$$

$$= \gamma_n(\{ x \in \mathbb{R}^n : \sqrt{\tfrac{n}{k}} \|Px\| \geq \tfrac{1}{1-\varepsilon} \|x\| \})$$

$$\leq e^{-\varepsilon^2 k/4} + e^{-\varepsilon^2 n/4},$$

Similarly, the result for $\sqrt{\frac{n}{k}} \|Py\| \leq (1-\varepsilon) \cancel{\|y\|}$ follows.

Given P_1, orthog. proj. onto V_1.

Want to find orth. proj. P_2 mapping

onto $\{ y \in \mathbb{R}^n : y = O x, \; x \in V_1 \} \overset{?}{=} O(V_1)$.

Given orthonorm. basis $\{ e_1, e_2, \dots, e_k \}$ of V_1

then $\{ O e_1, O e_2, \dots O e_k \}$

is orthon. basis of V_2.

Hence,
$$P_1 x = \sum_{j=1}^{k} \langle x, e_j \rangle e_j$$

and
$$P_2 x = \sum_{j=1}^{k} \underbrace{\langle x, O e_j \rangle}_{\langle O^* x, e_j \rangle} O e_j$$

$$= O \sum_{j=1}^{k} \langle \boxed{O^* x}, e_j \rangle e_j$$

$$= O P_1 O^* x$$

We recall the left-invariant Haar measure on $O(n)$, ν_n.

For fixed k-dim subspace V_1, assoc. orth. proj. P_{V_1}, we define a map

$$\Psi : O(n) \to G_k(\mathbb{R}^n), \quad O \mapsto O P_{V_1} O^* .$$

In terms of sets of subspaces in $G_k(\mathbb{R}^n)$, if $\mathcal{V}$ is an open set in $G_k(\mathbb{R}^n)$, let

$$\mu_{n,k}(\mathcal{V}) = \nu(\{ U \in O(n) : U(V_1) \in \mathcal{V} \}).$$

<u>Lemma</u> In above setting, with $x \in S^{n-1}$,

$$\mu_{n,k}\left(\left\{ V \in G_k(\mathbb{R}^n) : \sqrt{\tfrac{n}{k}} \|P_V x\| \geq \tfrac{1}{1-\varepsilon} \right\}\right) \leq e^{-\varepsilon^2 k/4} + e^{-\varepsilon^2 n/4}$$

and

$$\mu_{n,k}\left(\left\{ V \in G_k(\mathbb{R}^n) : \sqrt{\tfrac{n}{k}} \|P_V x\| \leq 1-\varepsilon \right\}\right) \leq e^{-\varepsilon^2 k/4} + e^{-\varepsilon^2 n/4} .$$

Pf. WLOG, choose $\|x\|=1$.

Choose any fixed k-dim subspace V_1 of $\mathbb{R}^n$,

and for $U \in O(n)$, let $V = U(V_1)$.

We use that $U \mapsto U(V_1)$ induces the

measure $\mu_{n,k}$ from Haar measure ν_n on $O(n)$.

This implies

$$\mu_{n,k}\left(\left\{ V \in G_k(\mathbb{R}^n) : \sqrt{\tfrac{n}{k}}\, \|P_V x\| \geq \frac{1}{1-\varepsilon} \right\}\right)$$

$$= \nu_n\left(\left\{ U \in O(n) : \sqrt{\tfrac{n}{k}}\, \|P_{U(V_1)} x\| \geq \frac{1}{1-\varepsilon} \right\}\right)$$

The projected length of the vector x is

$$\|P_{U(V_1)}\, x\| = \| U \boxed{P_{V_1} U^* x} \| .$$

$$= \| P_{V_1} \underbrace{U^* x} \| .$$

We note that the map ϕ_x,

$$\phi_x : O(n) \to S^{n-1}$$

$$U \mapsto U^* x$$

induces image measure μ_n on S^{n-1}.

Thus,

$$\nu_n \left(\left\{ U \in O(n) : \sqrt{\tfrac{n}{k}}\, \| P_{U(V_1)}\, x \| \geq \tfrac{1}{1-\varepsilon} \right\} \right)$$

$$= \mu_n \left(\left\{ y \in S^{n-1} : \sqrt{\tfrac{n}{k}}\, \| P_{V_1} y \| \geq \tfrac{1}{1-\varepsilon} \right\} \right) \leq e^{-\varepsilon^2 k/4} + e^{-\varepsilon^2 n/4} .$$

Similarly,

$$\nu_n\left(\left\{u \in O(n) : \sqrt{\tfrac{n}{k}}\, \| P_{u(V_1)} x \| \leq 1-\varepsilon \right\}\right)$$

$$= \mu_n\left(\left\{ y \in S^{n-1} : \sqrt{\tfrac{n}{k}}\, \| P_{V_1} y \| \leq 1-\varepsilon \right\}\right) \leq \ldots \qquad \square$$

Summary: Norm reduction for vectors under orth. proj. governed by measure $\mu_{n,k}$ is "often" given by factor $\sqrt{\tfrac{k}{n}}\,(1 \pm \varepsilon)$.

<u>Thm</u> (Johnson–Lindenstrauss, Dasgupta – Gupta)

Let $a_1, \ldots a_N$ be points in $\mathbb{R}^n$.

given $\varepsilon > 0$, choose $k \in \mathbb{N}$ s.th.

$$N(N-1)\left(e^{-k\varepsilon^2/4} + e^{-n\varepsilon^2/4}\right) \leq \frac{1}{3}$$

and $G_k(\mathbb{R}^n)$ is Grassmannian, with measure $\mu_{n,k}$

then

$$\mu_{n,k}\Big(\Big\{V \in G_k(\mathbb{R}^n) : \quad (1-\varepsilon)\,\|a_i - a_j\| \leq \sqrt{\tfrac{n}{k}}\,\|P_V(a_i - a_j)\| \leq \frac{1}{1-\varepsilon}\,\|a_i - a_j\|\Big\}; \text{ for all } 1 \leq i,j \leq N\Big\}\Big)$$

$$\geq \frac{2}{3}.$$