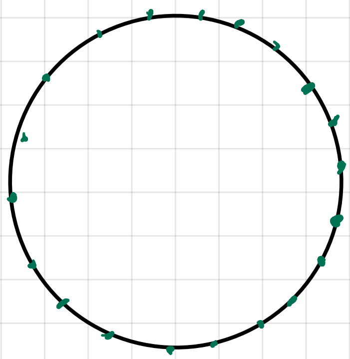
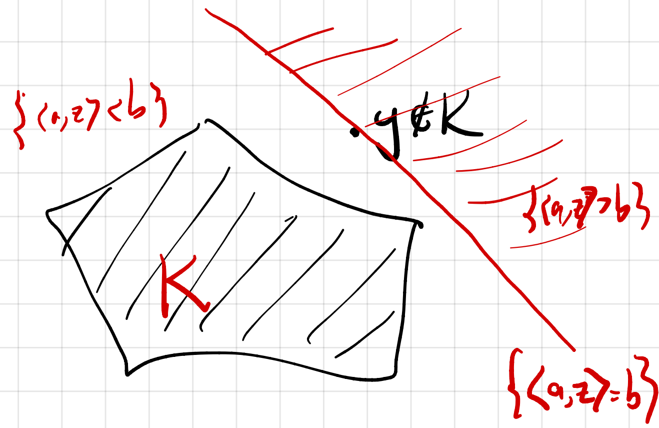
$\boxed{\text{25 Mar 2022}}$ Tail Bounds for Matrix Sums

Last time we constructed a set $C(d,\gamma)$ of $\leq \left(\frac{2e}{\gamma}\right)^d$ unit vectors in $\mathbb{R}^d$ such that $\forall x \in \mathbb{R}^d$, $\|x\|=1$, $\exists w \in C(d,\gamma)$ $\langle w, x\rangle > 1-\gamma$.





Lemma. For every $x \in \mathbb{R}^d$, $\|x\|=1$, the vector $(1-\gamma)x$ is a convex combination of vectors in $C(d,\gamma)$.

Proof. Let $K = \{$all cvx comb's of $C(d,\gamma)\}$.

If $(1-\gamma)x \notin K$ we will derive a contradiction.

Separating Hyperplane Principle says $\exists\, a \in \mathbb{R}^d$, $b \in \mathbb{R}$ s.t.

$$\langle a, (1-\gamma)x\rangle = b$$

$$\langle a, w\rangle < b \qquad \forall\, w \in C(d,\gamma)$$

By scaling a, and scaling b accordingly we can assume $\|a\|=1$.

$$\langle a, (1-\gamma)x\rangle = b \qquad \langle a,(1-\gamma)x\rangle = (1-\gamma)\langle a,x\rangle \leq 1-\gamma$$

$b \leq 1-\gamma$

$\forall w \in C(d,\gamma)$ $\langle a, w\rangle < b$ $\Longrightarrow$ $b > 1-\gamma.$

Recall $\exists\, w_0 \in C(d,\gamma)$ $\langle a, w_0\rangle > 1-\gamma.$

Lemma (Fancy Union Bound Lemma)

Suppose P is a random symmetric positive semidef matrix in $\mathbb{R}^{d\times d}$ and let $Q = \mathbb{E}[P]$.
Suppose Q is positive definite. Then $\forall\, 0<\beta<1$
$\forall\, 0<\gamma<1$

$$\Pr\left(\exists x \;\; \frac{\langle x, Px\rangle}{\langle x, Qx\rangle} \notin (1-\beta, 1+\beta)\right)$$

union of infinitely many events

$$\leq \left(\frac{2e}{\gamma}\right)^d \sup_{\substack{x\in\mathbb{R}^d \\ \|x\|=1}} \Pr\left(\frac{\langle x, Px\rangle}{\langle x, Qx\rangle} \notin \left(1-(1-\gamma)^2\beta,\, 1+(1-\gamma)^2\beta\right)\right)$$

infinitely many events, each with small probability

Cor. Suppose P and Q are as above, and suppose $\forall\, x\in\mathbb{R}^d$ $\|x\|=1$, $\Pr\left(\frac{\langle x, Px\rangle}{\langle x, Qx\rangle} \notin (1-\alpha, 1+\alpha)\right) < e^{-K\alpha^2 n}$

Then $\Pr\left(\exists x\in\mathbb{R}^d \;\; x\neq 0 \;\; \frac{\langle x, Px\rangle}{\langle x, Qx\rangle} \notin \left(1-(1-\gamma)^{-2}\alpha,\, 1+(1-\gamma)^{-2}\alpha\right)\right)$

$$< \exp\left(-K\alpha^2 n + d\ln\left(\frac{2e}{\gamma}\right)\right)$$

Application to SVD analysis.

Recall: A is a $d\times n$ matrix with columns $a_1, a_2, \ldots, a_n \in \mathbb{R}^d$ drawn independently from $N(0, BB^T)$.

Singular vector $\hat{v}$ is top eigenvect of AA^T.

v_1 is — — — — BB^T.

(Both normalized so $\|\hat{v}\| = \|v_1\| = 1$.)

Trying to prove $\langle \hat{v}, v_1 \rangle \geq 1-\varepsilon$ w. prob. $\geq 1-\delta$

if n is large enough.

Apply lemma above with $P = \frac{1}{n} AA^T$, $Q = BB^T$.

$$AA^T = \sum_{i=1}^{n} a_i a_i^T$$

each is equal to BB^T

$$\mathbb{E}[AA^T] = \sum_{i=1}^{n} \mathbb{E}[a_i a_i^T] = nBB^T,$$

For any $x \in \mathbb{R}^d$,

$$\langle x, Px \rangle = \sum_{i=1}^{n} \langle x, a_i a_i^T x \rangle$$

sum of indep, ident distrib. squares of Gaussians.

$$= \sum_{i=1}^{n} (x^T a_i)(a_i^T x) = \sum_{i=1}^{n} \langle a_i, x \rangle^2$$

<u>FACT.</u> (To be proven in typeset notes)

If $X_1, \dots, X_n$ are iid squares of mean-zero Gaussians

$$\Pr\left(\frac{\sum_{i=1}^{n} X_i}{\mathbb{E} \sum_{i=1}^{n} X_i} \notin (1-\alpha, 1+\alpha) \right) < \exp\left(-\frac{1}{6}\alpha^2 n\right).$$

Using Corollary above, we conclude

$$\Pr\left(\exists x \in \mathbb{R}^d \quad \frac{\langle x, Px\rangle}{\langle x, Qx\rangle} \notin \left(1 - \frac{\alpha}{(1-\gamma)^2}, 1 + \frac{\alpha}{(1-\gamma)^2}\right)\right)$$

$$< \exp\left(-\frac{1}{6}\alpha^2 n + d \ln\left(\frac{2e}{\gamma}\right)\right)$$

Choose n to make this < δ.

Choose α, γ s.t. when this event doesn't happen $\langle \hat{v}, v_1\rangle \geq 1-\varepsilon$.