MATH 7251: High-Dimensional Probability	**Spring 2022**

Lecture 2

Name: Tyler LaBonte **Professor:** Konstantin Tikhomorov

1 Linear Algebra Review Continued

1.1 Trace

Suppose $\mathbf{A}$ is an $n \times n$ matrix. Then the trace of $\mathbf{A}$ is $\operatorname{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$, the sum of the diagonal elements. A crucial property is that $\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$. This implies that if $\mathbf{A}$ is diagonalizable, that is, $\mathbf{A} = \mathbf{M}^{-1}\mathbf{DM}$ with $\mathbf{D}$ diagonal, then

$$\operatorname{tr}(\mathbf{A}) = \operatorname{tr}\left(\mathbf{M}^{-1}\mathbf{DM}\right) = \operatorname{tr}\left(\mathbf{DMM}^{-1}\right) = \operatorname{tr}(\mathbf{D}) = \sum_{i=1}^{n} \lambda_i(\mathbf{A}). \tag{1}$$

By continuity, this argument extends to all square matrices.

1.2 Hilbert-Schmidt Norm

The Hilbert-Schmidt norm of an $n \times m$ matrix $\mathbf{A}$ is $\|\mathbf{A}\|_{\mathrm{HS}} = \sqrt{\sum_{i,j} a_{ij}^2}$. Note that

$$\|\mathbf{A}\|_{\mathrm{HS}}^2 = \operatorname{tr}\left(\mathbf{A}^\top \mathbf{A}\right) = \operatorname{tr}\left(\mathbf{A}\mathbf{A}^\top\right) = \sum_{i=1}^{m} s_i(\mathbf{A})^2, \tag{2}$$

that is, the Hilbert-Schmidt norm is the ℓ_2 norm of the vector of singular values.

1.3 Projection

Suppose $\mathbf{P}$ is an $n \times n$ matrix. Then, $\mathbf{P}$ is a projection operator if $\mathbf{P}^2 = \mathbf{P}$ – that is, $\mathbf{P}$ acts trivially on its column space. If its null space is additionally orthogonal to its column space, $\mathbf{P}$ is called orthogonal.

Lemma 1.1

If $\mathbf{P}$ is a projection operator, then $\mathbf{P}$ is an orthogonal projection iff it is a symmetric matrix.

Note that an orthogonal projection and an orthogonal matrix are completely different notions. Suppose $\mathbf{P}$ is an orthogonal projection and E be its column space with $\dim(E) = \operatorname{rk}(\mathbf{P}) = k$. Let $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k\}$ be an orthonormal basis in E. Then we can write $\mathbf{P} = \sum_{i=1}^{k} \boldsymbol{x}_i \boldsymbol{x}_i^\top$. This means that the eigenvalues of $\mathbf{P}$ can only be zero or one, and the multiplicity of one is k.

Note also that if $\mathbf{P}$ is an orthogonal projection, then $\mathbf{I} - \mathbf{P}$ is an orthogonal projection on the complement.

2 Packing and Covering

If $V, W \subset \mathbb{R}^n$ are two sets, then the Minkowski sum of V, W is defined as $V + W = \{\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} \in V, \boldsymbol{y} \in W$. A set $K \subset \mathbb{R}^n$ is convex if for all $\boldsymbol{x}, \boldsymbol{y} \in K$ and $\lambda \in (0,1), \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in K$. A convex body is a compact convex set with a non-empty interior. Suppose that K is a convex body such that the origin is in the interior of K. Then, the Minkowski functional associated with K is

$$\|\boldsymbol{x}\|_K = \inf\{\lambda > 0 : \boldsymbol{x} \in \lambda K\}, \boldsymbol{x} \in \mathbb{R}^n. \tag{3}$$

if K is origin-symmetric ($K = -K$), then $\|\cdot\|_K$ is a norm in $\mathbb{R}^n$. So, we have a bijection between norms in $\mathbb{R}^n$ and origin-symmetric convex bodies. One special example is B_p^n, the closed unit ball of ℓ_p norm in $\mathbb{R}^n$. Here, $\|\cdot\|_{B_p^n} = \|\cdot\|_p$. Finally, suppose that (T, ρ) is a metric space.

Let $S \subset T$ and $\epsilon < 0$. Then a subset $N \subset S$ is an ϵ-net in S if for every $\boldsymbol{x} \in S$ there exists $\boldsymbol{y} \in N$ with $p(\boldsymbol{x}, \boldsymbol{y}) \leq \epsilon$. That is, N is sufficiently dense so that there is a point in N "close enough" to every point in S. A subset $N \subset S$ is a maximal ϵ-separated subset of S if for all $\boldsymbol{x}, \boldsymbol{y} \in N$ with $\boldsymbol{x} \neq \boldsymbol{y}$, we have $\rho(\boldsymbol{x}, \boldsymbol{y}) \geq \epsilon$ and for all $\boldsymbol{x} \in S$ there exists $\boldsymbol{y} \in N$ with $\rho(\boldsymbol{x}, \boldsymbol{y}) \leq \epsilon$. Thus, maximal ϵ-separated subsets can be viewed as "efficient" ϵ-nets.

Let T be a subset of $\mathbb{R}^n$ and K a convex body in $\mathbb{R}^n$. Then, the collection of parallel translations of K: $(\boldsymbol{x}_\alpha + K)_{\alpha \in \mathcal{A}}$ is a packing in T if (1) $\boldsymbol{x}_\alpha + K \subset T$ for all $\alpha \in \mathcal{A}$ and (2) $(\boldsymbol{x}_{\alpha_1} + \mathrm{Int}(K)) \cap (\boldsymbol{x}_{\alpha_2} + \mathrm{Int}(K)) = \emptyset$ whenever $\alpha_1 \neq \alpha_2$. A packing is maximal if there is no vector $\boldsymbol{y} \in \mathbb{R}^n$ such that $\{(\boldsymbol{x}_\alpha + K)_{\alpha \in \mathcal{A}}, \boldsymbol{y} + K\}$ is a packing in T.

Lemma 2.1

Suppose $T \subset \mathbb{R}^n$ and K is a convex body and $\{\boldsymbol{x}_i + K\}_{i=1}^M$ is a packing in T. Then

$$M \cdot \mathrm{Vol}(K) \leq \mathrm{Vol}(T). \tag{4}$$

This implies that any packing of parallel translations of K in T can have cardinality at most $\mathrm{Vol}(T)/\mathrm{Vol}(K)$. This is often attained when T admits a tiling by K.

Assume K is an origin-symmetric convex set and let $\{\boldsymbol{x}_i\}_{i=1}^M$ be a maximal ϵ-separated set in T wrt the Minkowski functional $\|\cdot\|_K$. Then, the collection $\{x_i + \frac{\epsilon}{2}K\}_{i=1}^M$ is a packing in $T + \frac{\epsilon}{2}K$. (We are transforming a set of points into a packing of convex bodies). This implies that the cardinality of any maximal ϵ-separated set in T is at most $\mathrm{Vol}(T + \frac{\epsilon}{2}K)/\mathrm{Vol}(\frac{\epsilon}{2}K)$.

As a corollary, if K is an origin symmetric convex body in $\mathbb{R}^n$ and $\epsilon > 0$, then K admits an ϵ-net wrt $\|\cdot\|_K$ of cardinality at most $(1 + \frac{2}{\epsilon})^n$. This means that K can be covered by at most $(1 + \frac{2}{\epsilon})^n$ parallel translations of K.