Optimization-Based Data Analysis – Brett Bernstein

Recitation 1

1. If a matrix $U \in \mathbb{R}^{n\times n}$ is orthogonal, must its transpose also be orthogonal?

   Solution. Yes. As U is orthogonal, we have $U^TU = I$. Thus U, U^T are inverses of each other, so $UU^T = I$. This shows $(U^T)^TU^T = I$ proving U^T is orthogonal.

2. Suppose $A \in \mathbb{R}^{m\times n}$ with $\text{trace}(AA^T) = 0$. What can be said about A?

   Solution. We must have $A = 0$. Note that

   $$\text{trace}(AA^T) = \langle A^T, A^T\rangle = \|A^T\|_F^2 = 0,$$

   implying $A^T = 0$.

3. Prove or disprove: If $A, B \in \mathbb{R}^{n\times n}$ with $\text{trace}(A) = 0 = \text{trace}(B)$ then $\text{trace}(AB) = 0$.

   Solution. False. Let
   $$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
   and note that $\text{trace}(A) = 0$ and $\text{trace}(A^2) = \text{trace}(I) = 2$.

4. Prove the converse to the Pythagorean theorem holds in a real inner product space: If $\|\vec{x}\|^2 + \|\vec{y}\|^2 = \|\vec{x} + \vec{y}\|^2$ then $\langle \vec{x}, \vec{y}\rangle = 0$.

   Solution. Note that

   $$\begin{aligned} \|\vec{x}+\vec{y}\|^2 = \langle \vec{x}+\vec{y}, \vec{x}+\vec{y}\rangle &= \langle \vec{x}, \vec{x}+\vec{y}\rangle + \langle \vec{y}, \vec{x}+\vec{y}\rangle && (1) \\ &= \langle \vec{x}, \vec{x}\rangle + \langle \vec{x}, \vec{y}\rangle + \langle \vec{y}, \vec{x}\rangle + \langle \vec{y}, \vec{y}\rangle && (2) \\ &= \|\vec{x}\|^2 + 2\langle \vec{x}, \vec{y}\rangle + \|\vec{y}\|^2. && (3) \end{aligned}$$

   By assumption, the last line is equal to $\|\vec{x}\|^2 + \|\vec{y}\|^2$ proving $\langle \vec{x}, \vec{y}\rangle = 0$.

5. Prove Bessel's inequality: Let $\vec{x}, \vec{u}_1, \ldots, \vec{u}_n \in V$ where V is a (real or complex) inner product space. Then
   $$\sum_{i=1}^n |\langle \vec{x}, \vec{u}_i\rangle|^2 \leq \|\vec{x}\|^2$$
   if $\vec{u}_1, \ldots, \vec{u}_n$ are orthonormal.

   Solution. Let $\mathcal{S} = \text{span}(\vec{u}_1, \ldots, \vec{u}_n)$. Then

   $$\mathcal{P}_{\mathcal{S}}\vec{x} = \sum_{i=1}^n \langle \vec{x}, \vec{u}_i\rangle \vec{u}_i.$$

By the Pythagorean theorem, we have

$$\|\mathcal{P}_{\mathcal{S}}\vec{x}\|^2 + \|\mathcal{P}_{\mathcal{S}^\perp}\vec{x}\|^2 = \|\vec{x}\|^2.$$

Thus

$$\|\vec{x}\|^2 \geq \|\mathcal{P}_{\mathcal{S}}\vec{x}\|^2 = \sum_{i=1}^{n} |\langle \vec{x}, \vec{u}_i \rangle|^2.$$

6. If V, W are subspaces of $\mathbb{R}^n$ with $\dim(V) + \dim(W) > n$ then there is a non-zero vector in $V \cap W$.

   Solution. Let $\mathcal{B}_V = (\vec{v}_1, \dots, \vec{v}_r)$ and $\mathcal{B}_W = (\vec{w}_1, \dots, \vec{w}_s)$ be bases for V, W, respectively. If $V \cap W = \{0\}$ then

   $$\mathcal{B} = (\vec{v}_1, \dots, \vec{v}_r, \vec{w}_1, \dots, \vec{w}_s)$$

   is linearly independent in $\mathbb{R}^n$ with $r + s > n$, a contradiction.

   To see that $\mathcal{B}$ is linearly independent, suppose

   $$a_1\vec{v}_1 + \cdots + a_r\vec{v}_r + b_1\vec{w}_1 + \cdots + b_s\vec{w}_s = 0,$$

   for some $a_1, \dots, a_r, b_1, \dots, b_s \in \mathbb{R}$. Then

   $$a_1\vec{v}_1 + \cdots + a_r\vec{v}_r = -b_1\vec{w}_1 - \cdots - b_s\vec{w}_s.$$

   Since $V \cap W = \{0\}$ by assumption, both sides must be zero. Since $\mathcal{B}_V, \mathcal{B}_W$ are linearly independent, we have $a_1 = \cdots = a_r = 0$ and $b_1 = \cdots = b_s = 0$.

   To see that $\mathbb{R}^n$ cannot have a linearly independent sequence $(\vec{x}_1, \dots, \vec{x}_k)$ with $k > n$, put $\vec{x}_1, \dots, \vec{x}_k$ as the columns of a matrix $A \in \mathbb{R}^{n \times k}$. Then $\operatorname{rank}(A) \leq n < k$ showing the columns cannot be linearly independent.

7. For any $\vec{x} \in \mathbb{R}^n$ show that

   $$\|\vec{x}\|_\infty \leq \|\vec{x}\|_2 \leq \|\vec{x}\|_1.$$

   Solution. For the first inequality, let m be such that $|\vec{x}[m]| = \|\vec{x}\|_\infty$. Then

   $$\|\vec{x}\|_2 = \sqrt{\sum_{i=1}^{n} \vec{x}[i]^2} \geq \sqrt{|\vec{x}[m]|^2} = |\vec{x}[m]| = \|\vec{x}\|_\infty.$$

   For the second inequality, note that

   $$\|\vec{x}\|_1^2 = \left(\sum_{i=1}^{n} |\vec{x}[i]|\right)^2 = \sum_{i=1}^{n} |\vec{x}[i]|^2 + \sum_{i \neq j} |\vec{x}[i]\vec{x}[j]| \geq \sum_{i=1}^{n} |\vec{x}[i]|^2 = \|\vec{x}\|_2^2.$$

   For an alternative proof of the second inequality, let $|\vec{x}| \in \mathbb{R}^n$ be the vector with $|\vec{x}|[i] = |\vec{x}[i]|$. Then

   $$\|\vec{x}\|_2^2 = |\vec{x}|^T|\vec{x}| = \operatorname{trace}(|\vec{x}|^T|\vec{x}|) = \operatorname{trace}(|\vec{x}||\vec{x}|^T) \leq \vec{1}^T|\vec{x}||\vec{x}|^T\vec{1} = \|\vec{x}\|_1^2,$$

   where $\vec{1} \in \mathbb{R}^n$ is the vector with 1 in every coordinate.

8. Consider the equation $A\vec{x} = \vec{b}$ where $A \in \mathbb{R}^{m\times n}$, $\vec{x} \in \mathbb{R}^n$, and $\vec{b} \in \mathbb{R}^m$.

   (a) Give conditions on $A, \vec{b}$ so that there is always an $\vec{x}$ satisfying the equation.

   (b) When is the solution unique?

   (c) Under what conditions does $A^T A\vec{x} = A^T\vec{b}$ have a solution?

   Solution.

   (a) $\vec{b} \in \mathrm{col}(A)$.

   (b) $\mathrm{null}(A) = \{0\}$, or equivalently, that A has full column rank.

   (c) It always has a solution. We will prove this by showing $\mathrm{col}(A^T A) = \mathrm{col}(A^T)$. As $A^T A\vec{x} = A^T(A\vec{x})$ we see that $\mathrm{col}(A^T A) \subseteq \mathrm{col}(A^T)$. To complete the proof we will show that $\dim(\mathrm{col}(A^T A)) = \dim(\mathrm{col}(A^T))$. We begin by showing $\mathrm{null}(A) = \mathrm{null}(A^T A)$. To see this note that $A\vec{x} = 0$ implies $A^T A\vec{x} = 0$ and $A^T A\vec{y} = 0$ implies
   $$0 = \vec{y}^T A^T A\vec{y} = \|A\vec{y}\|_2^2.$$
   This proves $\mathrm{null}(A) = \mathrm{null}(A^T A)$, which implies $\mathrm{rank}(A) = \mathrm{rank}(A^T A)$ since
   $$\mathrm{rank}(A) + \dim(\mathrm{null}(A)) = n = \mathrm{rank}(A^T A) + \dim(\mathrm{null}(A^T A)).$$
   Thus
   $$\dim(\mathrm{col}(A^T)) = \mathrm{rank}(A) = \mathrm{rank}(A^T A) = \dim(\mathrm{col}(A^T A)).$$

9. Let $f(\vec{x}) = A\vec{x}$ and $g(\vec{x}) = \vec{b}^T\vec{x}$ where $A \in \mathbb{R}^{m\times n}$ and $\vec{b} \in \mathbb{R}^n$. Give necessary and sufficient conditions for $\mathrm{null}(f) \subseteq \mathrm{null}(g)$.

   Solution. We will show that $\vec{b} \in \mathrm{row}(A)$ is a necessary and sufficient condition. If $\vec{b} \in \mathrm{row}(A)$ then we can write $\vec{b} = A^T\vec{y}$ for some $\vec{y} \in \mathbb{R}^m$. Thus if $A\vec{x} = 0$ then $\vec{b}^T\vec{x} = \vec{y}^T A\vec{x} = 0$. Conversely, suppose $\vec{b} \notin \mathrm{row}(A)$. Express $\vec{b}$ as
   $$\vec{b} = A^T\vec{y} + \vec{w},$$
   where $\vec{y} \in \mathbb{R}^m$ and $\vec{w} \in \mathrm{row}(A)^\perp$ with $\vec{w} \neq 0$. Then $A\vec{w} = 0$ but
   $$\vec{b}^T\vec{w} = (\vec{y}^T A + \vec{w}^T)\vec{w} = \|\vec{w}^T\|_2^2 > 0.$$